CORC Technical Report TR-2004-11
Robust Active Portfolio Management[*]

E. Erdoğan[†] D. Goldfarb[‡] G. Iyengar[§]

November 27, 2006

Abstract

In this paper we construct robust models for active portfolio management in a market with transaction costs. The goal of these robust models is to control the impact of estimation errors in the values of the market parameters on the performance of the portfolio strategy. Our models can handle a large class of piecewise convex transaction cost functions and allow one to impose additional side constraints such as bounds on the portfolio holdings, constraints on the portfolio beta, and limits on cash and industry exposure. We show that the optimal portfolios can be computed by solving second-order cone programs – a class of optimization problems with a worst case complexity (i.e., cost) that is comparable to that for solving convex quadratic programs (e.g. the Markowitz portfolio selection problem). We tested our robust strategies on simulated data and on real market data from 2000-2003 imposing realistic transaction costs. In these tests, the proposed robust active portfolio management strategies significantly outperformed the S&P 500 index without a significant increase in volatility.

1 Introduction

Portfolio management is concerned with allocating capital over a number assets to maximize a suitably defined measure of "return" and minimize "risk". Although the role of diversification in reducing risk has long been a part of financial folklore, Markowitz [26, 27] formulated the first mathematical model that explicitly modeled the risk-return trade-off. In the Markowitz model the "return" of a portfolio is defined to be the expected value of the random portfolio return and the 'risk" is quantified by the variance of the random portfolio return. Markowitz showed that, for a specified lower bound on the return , the optimal portfolio can be computed by solving a convex quadratic program.

---

[*] Research supported in part by NSF grants CCR-00-09972, DMS-01-04282, DMS 06-06712, ONR grant N000140310514 and DOE grant GE-FG01-92ER-25126.

[†] IEOR Department, Columbia University, New York, New York 10027. Email: `ee168@columbia.edu`.

[‡] IEOR Department, Columbia University, New York, NY 10027. Email:`goldfarb@columbia.edu`.

[§] IEOR Department, Columbia University, New York, New York 10027. Email: `garud@columbia.edu`.

Report Documentation Page	*Form Approved* *OMB No. 0704-0188*

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE **27 NOV 2006**	2. REPORT TYPE	3. DATES COVERED **00-00-2006 to 00-00-2006**

4. TITLE AND SUBTITLE **Robust Active Portfolio Management**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Columbia University ,Department of Computer Science,New York,NY,10027**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

In this paper we construct robust models for active portfolio management in a market with transaction costs. The goal of these robust models is to control the impact of estimation errors in the values of the market parameters on the performance of the portfolio strategy. Our models can handle a large class of piecewise convex transaction cost functions and allow one to impose additional side constraints such as bounds on the portfolio holdings, constraints on the portfolio beta, and limits on cash and industry exposure. We show that the optimal portfolios can be computed by solving second-order cone programs { a class of optimization problems with a worst case complexity (i.e., cost) that is comparable to that for solving convex quadratic programs (e.g. the Markowitz portfolio selection problem). We tested our robust strategies on simulated data and on real market data from 2000-2003 imposing realistic transaction costs. In these tests, the proposed robust active portfolio management strategies significantly outperformed the S&P 500 index without a significant increase in volatility.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**42**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

The analytical tractability of the Markowitz mean-variance model led to development of the Capital Asset Pricing Model (CAPM) for asset pricing [35, 29, 23] which remains one of the most widely used models for equilibrium asset prices. In spite of its theoretical success, the practical impact of the Markowitz model has been quite limited because the model often produces "error-maximized and investment irrelevant portfolios" [28]. This behavior is a manifestation of the fact that the Markowitz-optimal portfolio is extremely sensitive to estimation errors in the parameters, i.e. the mean and variance of asset returns, and often amplifies the errors several-fold. A number of methods have been proposed for mitigating the effects of parameter uncertainty, such as constraining the portfolio weights [8, 14] and using scenario or re-sampling based stochastic programming methods [28, 40], James-Stein "shrinkage" estimates [9] and other Bayesian estimation techniques [21, 13, 7, 39, 13, 32, 33, 22]. However, these approaches do not explicitly account for parameter uncertainty in the portfolio construction step.

To model the effect of data uncertainty in optimization problems, Ben-Tal and Nemirovski [4, 3, 5] introduced a deterministic framework called *robust optimization.* Using this approach, Goldfarb and Iyengar [15] proposed a robust portfolio selection model that assumes that the market parameters lie in known and bounded uncertainty sets, and computes a portfolio by solving a max-min mean-variance problem that assumes the worst case behavior of the parameters. The uncertainty sets in [15] correspond to confidence regions around point estimates of the parameters; consequently, there is a probabilistic guarantee on the robust portfolio's performance. Moreover, in [15] it was shown that for these uncertainty sets, the max-min problems can be reformulated as second-order cone programs (SOCPs). Since the computational complexity of an SOCP is comparable to that of a convex quadratic program of similar size and structure [31, 24], the computational effort required to compute a robust portfolio is comparable to that required to compute the Markowitz-optimal portfolio. (See [1, 24] for an introduction to SOCP. A MATLAB toolbox to solve SOCPs can be downloaded from [36].) In the computational experiments reported in [15], the performance of the robust portfolio was generally superior to that of the non-robust Markowitz portfolio. A robust asset allocation model with different uncertainty sets for the parameters has been proposed by Tütüncü and Koenig [38]

Unlike Bayesian approaches that obtain improved estimates of the mean and variance of asset returns by incorporating uncertainty into the underlying asset return models, robust optimization based approaches model the mean and variance to be uncertain within specified uncertainty sets. It is unclear whether either of these approaches is unequivocally superior in modeling parameter uncertainty. However, the robust optimization based approaches have significant computational advantage since the associated optimization problem scales very gracefully with the problem size and one can add many convex constraints such as bounds on the portfolio holdings, constraints on the portfolio beta, constraints on the transaction costs, limits on the "distance" of rebalanced portfolios, without affecting the complexity of solving the portfolio allocation problem. Furthermore, when the uncertainty sets are confidence regions around the maximum-likelihood estimates for the problem parameters, the robust approach provides probabilistic guarantees on the performance of the computed portfolios.

The goal in *active* portfolio management is to beat a given benchmark by using information that is not broadly available in the market. In this paper we show how to employ robust optimization techniques to significantly improve the performance of active portfolio management. The main contributions of this paper are as follows.

(a) We show how to incorporate robustness with respect to parameter perturbations into mean-variance models for active portfolio management. Active portfolio management attempts to outperform a given market index by carefully trading in assets that are priced incorrectly, i.e. purchasing undervalued assets and short-selling those that are over-valued (see [17] and §2 for details). Since errors in estimating the returns of assets are expected to have serious consequences for an active strategy, robust models are likely to result in portfolios with significantly superior performance. The results of our numerical experiments with both simulated and real market data clearly illustrate the benefits of using a robust model. The models proposed in [15] cannot be directly applied to active portfolio management because the performance measures relevant for active investing are always *relative* to the given index.

(b) We show how a very large class of piecewise convex trading cost functions can be incorporated into an active portfolio selection problem in a tractable manner. Since active portfolio strategies tend to execute many trades, properly modeling and managing trading costs are essential for the success of any practical active portfolio management model [25]. Our transaction cost models are motivated by a study by Loeb [25] where it is shown that transaction costs are a function of the amount traded and the market capitalization of the traded assets.

(c) We show that side-constraints such as net-zero alpha (see §3.1) and investor views can be easily incorporated into the robust mean-variance active portfolio selection problem.

(d) We propose alternative data-driven models for active portfolio management that minimize Value-at-Risk and Conditional-Value-at-Risk with respect to a given benchmark. We also show how to incorporate robustness with respect to both parametric and non-parametric perturbations into these data-driven models.

The organization of this paper is as follows. In §2 we introduce a robust mean-variance model for active portfolio management. In this section we discuss a robust factor model for the asset returns (see §2.1), a piece-wise convex trading cost function that can be calibrated from data (see §2.3) and show that the associated robust portfolio selection problem can be reformulated as an SOCP (see §2.4). In §3 we discuss how to incorporate side constraints such as net-zero alpha and investor views into a robust mean-variance active portfolio selection. In this section, we also discuss data-driven non-parametric models for active portfolio management. In § 4 we report the results of our computational experiments with simulated and real market data and in § 5 we conclude with a brief discussion of avenues for future research.

2 Robust Mean-Variance Active Portfolio Management

In this section we propose a robust mean-variance model for active portfolio management. Our model uses historical returns and equilibrium expected returns predicted by the CAPM to identify assets that are incorrectly priced in the market.

There is a fundamental inconsistency between the CAPM and active portfolio management. The CAPM assumes that markets are efficient, i.e. if one were to project the random return r_a of any asset on the random return r_b of the benchmark to obtain linear relation

$$\mathbf{E}[r_a] = \alpha_a + \beta_a \mathbf{E}[r_b],$$

the *expected exceptional return* α_a is identically equal to zero. In contrast, active management is predicated on the assumption that there exist assets for which $\alpha_a \neq 0$ and that these exceptional returns can be predicted. There is an ongoing debate on the validity of the efficient market hypothesis and the CAPM (see [12] and the references therein) that is beyond the scope of this paper. In this paper we take a data driven approach and investigate whether historical returns can be used to outperform a given index.

Active portfolio management consists of two steps. The first step is to forecast the vector of exceptional returns $\boldsymbol{\alpha}$. We are not aware of any statistically justifiable model for estimating $\boldsymbol{\alpha}$. Practitioners appear to employ an ad-hoc combination of techniques, such as asset pricing models, security and/or sector analysis and analyst views, to arrive at an estimate of $\boldsymbol{\alpha}$ [17]. We propose combining historical returns with the CAPM to estimate $\boldsymbol{\alpha}$. The second step is to use the estimates of $\boldsymbol{\alpha}$ to compute a portfolio $\boldsymbol{\phi}$ that outperforms the market. Treynor and Black [37] introduced a mean-variance model for computing such a portfolio. This model shifts the market portfolio toward (away from) the stocks that have positive (negative) "alpha". Black and Litterman [7, 19] refined this model by introducing uncertainty in the market parameters. Cvitanic et. al. [10] introduced a dynamic model that also considers hedging demands. In this paper, we propose using a robust maximum information ratio to manage an active portfolio.

2.1 A robust factor model for residual returns

We use the robust factor model proposed by Goldfarb and Iyengar [15] for modeling the residual returns. We begin by projecting the random return vector $\mathbf{r} \in \mathbb{R}^n_+$ of the n assets in the market on the benchmark return r_b to obtain the linear relation

$$\mathbf{r} = \boldsymbol{\beta} r_b + \Delta\mathbf{r}, \tag{1}$$

where $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_n)^T$,

$$\beta_i^b = \frac{\mathbf{cov}(r_i, r_b)}{\mathbf{var}(r_b)}, \tag{2}$$

denotes the "beta" of asset i, and $\Delta\mathbf{r}$ denotes the vector of residual or exceptional returns. In practice $\boldsymbol{\beta}$ is determined by applying a linear regression to daily asset and benchmark returns $\mathbf{r}^{(t)}$ and $r_b^{(t)}$ over a specified historical period (i.e., sequence of trading days, $t = 1, \ldots, T$).

We assume that the benchmark return $r_b \sim \mathcal{N}(\mu_b, \sigma_b^2)$ and the random residual return $\Delta\mathbf{r}$ is given by the factor model

$$\Delta\mathbf{r} = \boldsymbol{\alpha} + \mathbf{V}'\mathbf{f} + \boldsymbol{\epsilon}, \tag{3}$$

where $\boldsymbol{\alpha} = \mathbf{E}[\Delta\mathbf{r}]$ denotes the vector of expected residual returns, $\mathbf{f} \sim \mathcal{N}(\mathbf{0}, \mathbf{F}) \in \mathbb{R}^m$ $(m << n)$ denotes the returns on the m factors driving the market, $\mathbf{V} \in \mathbb{R}^{m \times n}$ denotes the factor loading matrix of the n assets, and $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{D}) \in \mathbb{R}^n$ is independent of $\mathbf{f}$. The notation $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$ denotes that $\mathbf{x}$ is a multivariate Normal random variable with mean vector $\boldsymbol{\mu}$ and covariance matrix Σ. The covariance matrices $\mathbf{F} \succ \mathbf{0}$ and $\mathbf{D} = \mathbf{diag}(\mathbf{d}) \succ \mathbf{0}$. (The notation $\mathbf{A} \succ \mathbf{0}$ denotes that $\mathbf{A}$ is positive definite.) Thus, $\Delta\mathbf{r} \sim \mathcal{N}(\boldsymbol{\alpha}, \mathbf{V}'\mathbf{F}\mathbf{V} + \mathbf{D})$.

In practice the parameters $\boldsymbol{\alpha}$, $\mathbf{F}$, $\mathbf{V}$, and $\mathbf{D}$ in the factor model (3) are estimated from limited historical data and are, therefore, subject to estimation noise. To deal with these estimation errors, we assume that the expected residual return $\boldsymbol{\alpha}$, the factor loading matrix $\mathbf{V}$ and the covariance matrix $\mathbf{D}$ of the residual returns are uncertain, with uncertainty sets given by the confidence regions around the maximum-likelihood estimates of $\boldsymbol{\alpha}$, $\mathbf{V}$ and $\mathbf{D}$, respectively. Thus, the individual diagonal elements d_i of $\mathbf{D}$ lie in an interval $[\overline{d}_i - \delta_i, \overline{d}_i + \delta_i]$, i.e. $\mathbf{D}$ belongs to the uncertainty set $\mathcal{S}_d$ given by

$$\mathcal{S}_d = \left\{\mathbf{D} : \mathbf{D} = \mathbf{diag}(\mathbf{d}),\ d_i = \overline{d}_i + \Delta d_i,\ |\Delta d_i| \le \delta_i,\ i = 1, ..., n\right\}, \tag{4}$$

and the uncertainty set $\mathcal{S}_v$ for the factor loading matrix $\mathbf{V}$ is of the form

$$\mathcal{S}_v = \left\{\mathbf{V} : \mathbf{V} = \mathbf{V}_0 + \mathbf{W},\ \|\mathbf{W}_i\|_g \le \rho_i,\ i = 1, ..., n\right\}, \tag{5}$$

where $\mathbf{W}_i$ denotes the $i-$th column of $\mathbf{W}$ and $\|\mathbf{w}\|_g = \sqrt{\mathbf{w}'\mathbf{G}\mathbf{w}}$ denotes the elliptic norm of $\mathbf{w}$ with respect to a symmetric positive definite matrix $\mathbf{G}$, and the uncertainty set for the expected residual return vector $\boldsymbol{\alpha}$ is of the form

$$\boldsymbol{\alpha} \in S_\alpha = \left\{\boldsymbol{\alpha} : \boldsymbol{\alpha} = \boldsymbol{\alpha}_0 + \boldsymbol{\xi}, |\xi_i| \le \eta_i, i = 1, ..., n\right\}. \tag{6}$$

Methods for computing $\boldsymbol{\alpha}_0$, $\boldsymbol{\eta}$, $\mathbf{V}_0$, $\mathbf{G}$, $\boldsymbol{\rho}$, $\overline{\mathbf{d}}$ and $\boldsymbol{\delta}$ from historical data are discussed in Appendix A. For a detailed justification for using confidence regions as uncertainty sets, see [15]. For simplicity, we assume that the factor covariance matrix $\mathbf{F}$ is certain but this assumption can also be relaxed and all of the methods proposed in this paper can be extended to this case; see [15] for details.

2.2 A robust portfolio selection model with transaction costs

Suppose $\bar{\boldsymbol{\phi}} \in \mathbb{R}^n$ denotes the current portfolio, i.e. $\bar{\phi}_i$ is the amount in dollars invested in asset i. In this section we propose an optimization model that computes a new re-balanced portfolio $\boldsymbol{\phi}$.

Let $\mathbf{z}, \mathbf{y} \geq \mathbf{0}$ denote, respectively, the dollar amounts of assets bought or sold. Then the resulting rebalanced portfolio $\boldsymbol{\phi} \in \mathbb{R}^n$ is given by

$$\boldsymbol{\phi} = \bar{\boldsymbol{\phi}} + \mathbf{z} - \mathbf{y}. \tag{7}$$

Let $T(\mathbf{z}, \mathbf{y})$ denote the cost incurred by the trade $(\mathbf{z}, \mathbf{y})$. Then we must have

$$\mathbf{1}^T \boldsymbol{\phi} + T(\mathbf{z}, \mathbf{y}) = \mathbf{1}^T \bar{\boldsymbol{\phi}}, \tag{8}$$

where $\mathbf{1} \in \mathbb{R}^n$ denotes a vector with every component equal to 1. From (7) and (8) it follows that $T(\mathbf{z}, \mathbf{y}) = \mathbf{1}'(\mathbf{y} - \mathbf{z})$. Note that we have ignored consumption and fresh inflow of capital. This can be easily modeled by defining a suitable consumption function.

We assume that the transaction cost function $T(\mathbf{z}, \mathbf{y})$ is continuous, increasing, and convex in $(\mathbf{z}, \mathbf{y})$. (In § 2.3 we discuss a particular specification of the transaction cost function.) In order to control the amount of the transaction costs incurred we impose the constraint,

$$T(\mathbf{z}, \mathbf{y}) = \mathbf{1}'(\mathbf{y} - \mathbf{z}) \leq \theta w, \tag{9}$$

where $\theta \in [0, 1]$ is a parameter, and w is the amount of the total wealth invested in the rebalanced portfolio $\boldsymbol{\phi}$, i.e. the wealth *after* the trade has been executed and the transaction costs are paid. In addition, we impose upper and lower bounds on the portfolio holdings as a percentage of the total wealth, i.e.

$$-w\mathbf{v} \leq \boldsymbol{\phi} \leq w\mathbf{u} \tag{10}$$

A positive lower bound $\mathbf{v} > 0$ implies that short sales are allowed.

Next, we select a suitable objective for the portfolio re-balancing step. The goal of active portfolio management is to outperform an index by investing in assets are mispriced with respect to the benchmark $\boldsymbol{\phi}_b$, i.e. assets for which the expected excess return $\alpha_i \neq \mathbf{0}$. When the excess return $\Delta\mathbf{r} \equiv \mathbf{0}$ the active manager would hold $\boldsymbol{\phi} = w\boldsymbol{\phi}_b$, where $\boldsymbol{\phi}_b$ is the normalized ($\mathbf{1}'\boldsymbol{\phi}_b = 1$) index portfolio. On the other hand, when $\Delta\mathbf{r} \neq \mathbf{0}$ we expect an active portfolio manage to hold a portfolio $\boldsymbol{\phi} \neq w\boldsymbol{\phi}_b$. Let

$$\boldsymbol{\psi} = \boldsymbol{\phi} - w\boldsymbol{\phi}_b, \tag{11}$$

denote the *active* component $\boldsymbol{\psi}$ of the portfolio $\boldsymbol{\phi}$, i.e. $\boldsymbol{\psi}$ is the difference between the portfolio $\boldsymbol{\phi}$ and the *passive* portfolio $w\boldsymbol{\phi}_b$ which invests the entire net wealth w in the market portfolio $\boldsymbol{\phi}_b$. The random return $\mathbf{r}'\boldsymbol{\phi}$ of the portfolio $\boldsymbol{\phi}$ can be decomposed as follows

$$\mathbf{r}'\boldsymbol{\phi} = (\Delta\mathbf{r} + r_b\boldsymbol{\beta})'\boldsymbol{\phi} = \Delta\mathbf{r}'\boldsymbol{\phi} + (\boldsymbol{\beta}'\boldsymbol{\phi})r_b = \Delta\mathbf{r}'\boldsymbol{\phi} + (\boldsymbol{\beta}'\boldsymbol{\psi})r_b + wr_b(\boldsymbol{\beta}'\boldsymbol{\phi}_b) = \Delta\mathbf{r}'\boldsymbol{\phi} + (\boldsymbol{\beta}'\boldsymbol{\psi})r_b + wr_b, \tag{12}$$

where we have used (1) and (11) and the fact that $\Delta\mathbf{r}'\boldsymbol{\phi}_b = (\mathbf{r} - r_b\boldsymbol{\beta})'\boldsymbol{\phi}_b = r_b - r_b\boldsymbol{\beta}'\boldsymbol{\phi}_b \equiv 0$. In the active portfolio selection literature, the component $(\boldsymbol{\beta}'\boldsymbol{\psi})r_b$ that is perfectly correlated with the random benchmark return r_b is called the return from *benchmark timing* [17], i.e. the return derived by carefully *timing* the purchase/sale of the benchmark. Since such timing is likely to be very sensitive to market data, the term $\boldsymbol{\beta}'\boldsymbol{\psi}$ is very often negative. To protect against this, most active portfolio selection strategies choose $\boldsymbol{\psi}$ to be uncorrelated with the benchmark return r_b, i.e. $\boldsymbol{\beta}'\boldsymbol{\psi} = 0$ or equivalently

$$\boldsymbol{\beta}'\boldsymbol{\phi} = w. \tag{13}$$

Since $\Delta\mathbf{r}$ is assumed to be a multivariate Normal vector, (12) and (13) imply that the probability that the active return falls short of the benchmark return r_b is given by

$$\mathbb{P}\left(\mathbf{r}'\boldsymbol{\phi} \leq wr_b\right) = \mathbb{P}(\Delta\mathbf{r}'\boldsymbol{\phi} \leq 0) = 1 - \Phi\big(\mathrm{IR}(\boldsymbol{\phi})\big), \tag{14}$$

where $\Phi(\cdot)$ denotes the cumulative density function (CDF) of a univariate standard Normal random variable and

$$\mathrm{IR}(\boldsymbol{\phi}) = \frac{\mathbf{E}[\Delta\mathbf{r}'\boldsymbol{\phi}]}{\sqrt{\mathbf{var}[\Delta\mathbf{r}'\boldsymbol{\phi}]}} = \frac{\boldsymbol{\alpha}'\boldsymbol{\phi}}{\sqrt{\boldsymbol{\phi}'(\mathbf{V}'\mathbf{F}\mathbf{V} + \mathbf{D})\boldsymbol{\phi}}}. \tag{15}$$

denotes the *information ratio* of the portfolio $\boldsymbol{\phi}$ [17]. The information $\mathrm{IR}(\boldsymbol{\phi})$ is the Sharpe Ratio of the exceptional returns of the portfolio $\boldsymbol{\phi}$.

If the parameters α, $\mathbf{V}$, $\mathbf{F}$ and $\mathbf{D}$ are known perfectly, then a natural objective for an active portfolio manager is to maximize

$$v_o(\boldsymbol{\phi}) = \mathrm{IR}(\boldsymbol{\phi}), \tag{16}$$

as first suggested by Treynor and Black [37] since this would minimize the probability of shortfall.

When the parameters are uncertain, a natural objective $v_r(\boldsymbol{\phi})$ is to maximize the worst case $\mathrm{IR}(\boldsymbol{\phi})$, i.e.

$$v_r(\boldsymbol{\phi}) \quad = \quad \min_{\{\boldsymbol{\alpha} \in S_\alpha, \mathbf{V}_\alpha \in S_v, \mathbf{D}_\alpha \in S_d\}} \left\{ \frac{\boldsymbol{\alpha}'\boldsymbol{\phi}}{\sqrt{\boldsymbol{\phi}'(\mathbf{V}'\mathbf{F}\mathbf{V} + \mathbf{D})\boldsymbol{\phi}}} \right\} \tag{17}$$

Combining (7), (9), and (10), with a relaxed version of (8), we have that the set of feasible $\boldsymbol{\phi} \in \mathbb{R}^n$ is given by

$$\Phi = \left\{ \boldsymbol{\phi} : \begin{array}{l} w - \mathbf{1}'\boldsymbol{\phi} = 0, \quad w - \boldsymbol{\beta}'\boldsymbol{\phi} = 0 \quad \boldsymbol{\phi} + \mathbf{y} - \mathbf{z} = \bar{\boldsymbol{\phi}}, \quad \mathbf{1}'(\mathbf{y} - \mathbf{z}) \leq \theta w, \\ \mathbf{1}'\boldsymbol{\phi} + T(\mathbf{z}, \mathbf{y}) \leq \mathbf{1}'\bar{\boldsymbol{\phi}}, \\ -w\mathbf{v} \leq \boldsymbol{\phi} \leq w\mathbf{u}, \qquad \mathbf{z},\ \mathbf{y},\ w\ \geq 0. \end{array} \right\} \tag{18}$$

That is, there exist $\mathbf{z}, \mathbf{y} \in \mathbb{R}^n_+$ and $w \in \mathbb{R}_+$ in addition to $\boldsymbol{\phi} \in \mathbb{R}^n$ that satisfy the constraints in (18). Note that by relaxing the constraint $\mathbf{1}'\boldsymbol{\phi} + T(\mathbf{z}, \mathbf{y}) = \mathbf{1}'\bar{\boldsymbol{\phi}}$, we ensure that Φ is convex. In Appendix B we show that this relaxation does not result in any loss of generality.

Thus, a non-robust optimal active portfolio is any optimal solution of the optimization problem $\max\{v_o(\boldsymbol{\phi}) : \boldsymbol{\phi} \in \Phi\}$. Since the objective $v_o(\boldsymbol{\phi})$ of this problem has the form of a Sharpe ratio, the problem is often referred to as the mean-variance active portfolio selection problem. A robust optimal active portfolio is any optimal solution of the optimization problem $\max\{v_r(\boldsymbol{\phi}) : \boldsymbol{\phi} \in \Phi\}$.

2.3 Transaction cost model

In this section, we discuss a specific form for $T(\mathbf{z}.\mathbf{y})$ that is quite general and yet allows one to solve the active portfolio selection problem very efficiently. Following Loeb [25] we assume that the transaction cost for each asset is primarily due to the effect that trade have on the price of the asset and hence, depends on the market capitalization of the asset and the size of the trade. We are implicitly assuming that many of the trades of the active portfolio are fairly large.

We begin with a relatively simple transaction cost function. We assume that $T(\mathbf{z}, \mathbf{y})$ is separable along assets, i.e.

$$T(\mathbf{z}, \mathbf{y}) = \sum_{i=1}^{n} T_i(z_i, y_i) \tag{19}$$

where $T_i(z_i, y_i),\ i = 1, \ldots, n$ is the transaction cost function for each asset i, and that the cost of buying and selling an asset $i,\ i = 1, \ldots, n$ is the same. Hence, $T_i(z_i, y_i) = T_i(z_i + y_i)$, i.e. $T_i(z_i, y_i)$ is a function of $z_i + y_i$. Furthermore, we assume that the function

$$T_i(x) = \max\{\vartheta_{i1} x, \vartheta_{i2} x^{\frac{3}{2}}\}, \tag{20}$$

where $\theta_{i1}, \theta_{i2} > 0$. Thus, $T_i(x)$ is a continuous, increasing, and piecewise convex function with two pieces:

$$T_i(z_i, y_i) = \begin{cases} \vartheta_{i1}(z_i + y_i), & \text{if } z_i + y_i \leq \pi_i \\ \vartheta_{i2}(z_i + y_i)^{\frac{3}{2}}, & \text{if } z_i + y_i > \pi_i \end{cases} \tag{21}$$

where the breakpoint π_i satisfies $\vartheta_{i2} = \vartheta_{i1} \pi_i^{-\frac{1}{2}}$.

The specific form in (21) is motivated by the data reported in [25]. Figure 1 displays the per unit transaction cost, i.e. $t(x) = \frac{T(x)}{x}$ as a function of the trading block size x for the three different market capitalizations classes: \$0.5–\$1bn, \$1–\$1.5bn, and > \$1.5bn. In order to better understand the structure of the function $t(x)$, in Figure 2 we plot $\ln(t(x))$ as a function $\ln(x)$. It is clear from this plot that for each of the capitalization classes, a piecewise linear function of $\ln(x)$ of the form

$$\ln(t(x)) \quad = \quad \begin{cases} a_1 & \text{if} \quad 0 \leq x \leq \pi \\ a_2 + a_3 \ln(x), & \text{if} \quad \pi \leq x \end{cases} \tag{22}$$

where a_1, a_2, a_3 and π are parameters, provides a very good fit for $\ln(t(x))$. The least squares estimates of the parameters (a_1, a_2, a_3, π) for the three asset classes are given in Table 1. (The best-fit curves are also plotted in Figure 1 and Figure 2.) The parameters in Table 1 imply that for assets with a market capitalization greater than \$1.5bn, the transaction cost function

$$T(x) = t(x)x = \begin{cases} 0.0115x, & x \leq e^{5.3647} = 213.7271, \\ 0.0012x^{1.4140}, & x \geq 213.7271. \end{cases}$$

Therefore, the function in (21) is a good approximation for transaction cost function implied by the best-fit curve. While the data in [25] is from 1982, it is reasonable to assume that the structure of transaction cost (i.e., continuous, piecewise, and convex) illustrated by the data would be

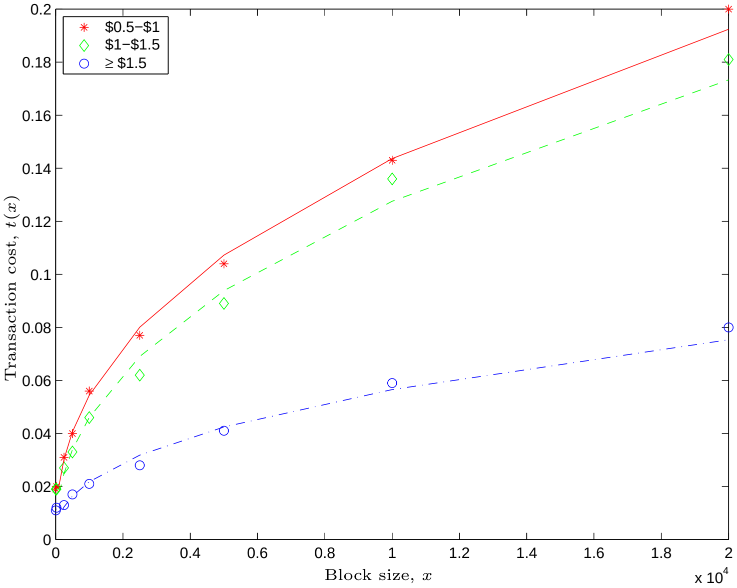

Figure 1: Per-unit transaction cost $t(x)$ as a function of market capitalization and the amount traded.

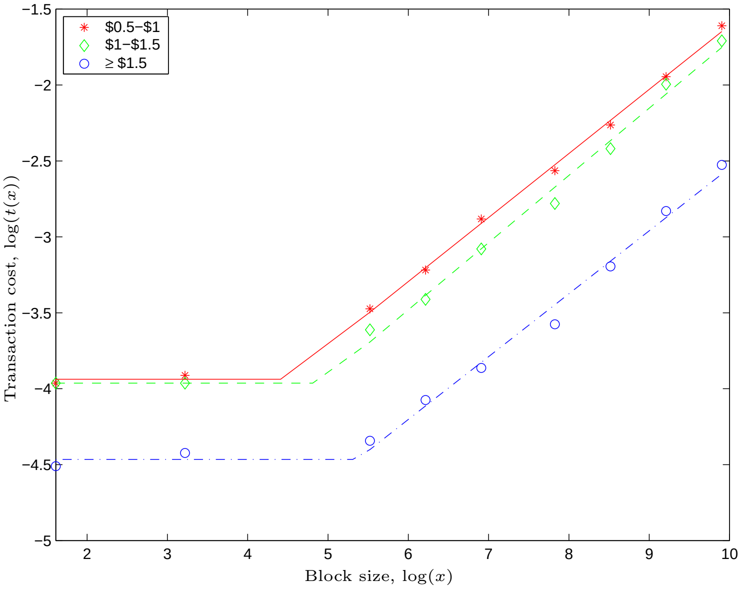

Figure 2: Logarithm of per-unit transaction cost $t(x)$ as a function of the logarithm of the block size x.

	\$0.5–\$1bn	\$1–\$1.5bn	$\geq$ \$1.5bn
$\ln(\pi)$	4.474	4.9087	5.3647
a_1	-3.9374	-3.9634	-4.4654
a_2	-5.8241	-6.1360	-6.6864
a_3	0.4217	0.4426	0.4140

Table 1: Coefficients of the least squares fits in Figure 2

appropriate to other time periods, although scales and numerical values might change. We discuss more general transaction cost functions later in this section.

Next we show that for transaction cost functions of the form (21), the set $\{\mathbf{z} \in \mathbf{R}^n_+, \mathbf{y} \in \mathbf{R}^n_+, \tau \in \mathbf{R}_+ | \mathbf{T}(\mathbf{z}, \mathbf{y}) \leq \tau\}$ can be represented as a collection of second-order cone (SOC) constraints of the form

$$\|\mathbf{A}\mathbf{x} + \mathbf{b}\| \leq \mathbf{c}^T\mathbf{x} + d,$$

where $\mathbf{x} \in \mathbb{R}^p$ is the decision variable, the norm $\|\cdot\|$ is the Euclidean norm, and all other quantities are constants.

From (19) and (20), it follows that

$$\begin{aligned} T(\mathbf{z},\mathbf{y}) \leq \tau \quad &\Leftrightarrow \quad \sum_{i=1}^n \tau_i \leq \tau, \quad T_i(z_i, y_i) = \max\{\vartheta_{i1}(z_i + y_i), \vartheta_{i2}(z_i + y_i)^{\frac{3}{2}}\} \leq \tau_i, \quad i = 1, \ldots, n \\ &\Leftrightarrow \quad \sum_{i=1}^n \tau_i \leq \tau \quad \vartheta_{i1}(z_i + y_i) \leq \tau_i, \quad \vartheta_{i2}(z_i + y_i)^{\frac{3}{2}} \leq \tau_i, \quad i = 1, \ldots, n. \end{aligned} \tag{23}$$

The following two results (e.g., see ([5]) or ([1]) show that a constraint of the form $\vartheta_{i2}(z_i + y_i)^{\frac{3}{2}} \leq \tau_i$ is equivalent to a collection of the SOC constraints.

Lemma 1 *The constraint* $\sqrt{x_1 x_2} \geq x_3$ *where* $x_1 \geq 0, x_2 \geq 0, x_3 \geq 0$ *is equivalent to the SOC constraint*

$$\left\| \begin{matrix} 2x_3 \\ x_1 - x_2 \end{matrix} \right\| \leq \quad x_1 + x_2, \quad x_1 \geq 0, x_2 \geq 0, x_3 \geq 0.$$

Proof: Squaring both sides of the norm inequality and simplifying yields the equivalence. ■

Lemma 2 *The set of constraints* $\vartheta x^{\frac{3}{2}} \leq t, x \geq 0, t \geq 0$, *where* $\vartheta \geq 0$ *is a constant, hold if and only if there exists* $t_1 \geq 0$ *satisfying*

$$\sqrt{\vartheta} x \leq \sqrt{t_1 t}, \quad t_1 \leq \sqrt{x}. \tag{24}$$

Proof: Suppose that $\vartheta x^{\frac{3}{2}} \leq t$. Then $t_1 = \sqrt{x}$ satisfies (24). On the other hand, suppose x, t, and t_1 satisfy (24). Then we have

$$\vartheta x^2 \leq t_1 t \leq x^{\frac{1}{2}} t \quad \Rightarrow \quad \vartheta x^{\frac{3}{2}} \leq t.$$

■

Lemmas 2 and 1 together with (23) imply that

$$\Phi = \left\{ \boldsymbol{\phi} : \begin{array}{l} w - \mathbf{1}'\boldsymbol{\phi} = 0, \quad w - \boldsymbol{\beta}'\boldsymbol{\phi} = 0 \quad \boldsymbol{\phi} + \mathbf{y} - \mathbf{z} = \bar{\boldsymbol{\phi}}, \quad \mathbf{1}'(\mathbf{y} - \mathbf{z}) \leq \theta w, \\ \mathbf{1}'\boldsymbol{\phi} + \sum_{i=1}^{n} \tau_i \leq \mathbf{1}'\bar{\boldsymbol{\phi}}, \\ \vartheta_{i1}(z_i + y_i) \leq \tau_i, \quad i = 1, \ldots, n \\ \left\| \begin{matrix} 2\sqrt{\vartheta_i}(z_i + y_i) \\ \tau_i - \kappa_i \end{matrix} \right\| \leq \tau_i + \kappa_i, \quad i = 1, \ldots, n, \\ \left\| \begin{matrix} 2\kappa_i \\ z_i + y_i - 1 \end{matrix} \right\| \leq z_i + y_i + 1, \quad i = 1, \ldots, n, \\ -w\mathbf{v} \leq \boldsymbol{\phi} \leq w\mathbf{u}, \qquad \mathbf{z},\ \mathbf{y},\ \boldsymbol{\kappa},\ \boldsymbol{\tau},\ w\ \geq 0. \end{array} \right\}$$

This reformulation has important practical implications since large-scale SOCPs can be solved very efficiently.

We now present a more general class of transaction cost functions that still allow the portfolio selection problem to be reformulated as an SOCP. We still assume that the transaction costs are separable by assets; however, we now allow for the transaction costs of buying and selling an asset i to be different; i.e.

$$T(\mathbf{z}, \mathbf{y}) = \sum_{i=1}^{n} T_i^{+}(z_i) + T_i^{-}(y_i).$$

All the transaction costs are assumed to have the following general form:

$$T(x) = \max_{1 \leq j \leq J} \Big\{ \sum_{k=1}^{K_j} a_{jk} x^{\rho_{jk}} \Big\}. \tag{25}$$

where for all j, k, $a_{jk} \geq 0$, and $\rho_{jk} \geq 1$ and rational. The transaction cost function in (25) is fairly general, e.g. functions of the form

$$T(x) = \begin{cases} a_{11}x, & x \leq b_1, \\ a_{21}x + a_{22}x^{\frac{4}{3}}, & b_1 \leq x \leq b_2, \\ a_{31}x^2, & x \geq b_2, \end{cases}$$

belong to the allowed class of transaction cost functions. However, it does not cover the class of all piecewise polynomial convex functions.

The constraint $T(x) \leq \tau$ if, and only if, there exist variables τ_{jk}, $k = 1, \ldots, K_j$, $j = 1, \ldots J$, such that

$$a_{jk} x^{\rho_{jk}} \leq \tau_{jk}, \qquad \sum_{k=1}^{K_j} \tau_{jk} \leq \tau \quad j = 1, \ldots, J.$$

The following result shows that constraints of the form $\theta x^{\frac{a}{b}} \leq t$ for $a \geq b \geq 1$ and integer can be reformulated as a collection of second-order cone constraints.

Lemma 3 ([5]) *Let a and b be positive integers with $a > b$. The set of constraints $\vartheta x^{\frac{a}{b}} \le t, x \ge 0, t \ge 0$, where $\vartheta \ge 0$ is a constant, hold if and only if there exists nonnegative variables t_{kl}, $k = 1, \ldots, q$, $l = 1, \ldots, 2^{q+1-k}$ where q is the smallest positive integer such that $a \le 2^q$, satisfying*

$$\begin{array}{rcl}
\vartheta^{b/2^q} x & \le & \sqrt{t_{q1} t_{q2}}, \\
t_{kl} & \le & \sqrt{t_{k-1,2l-1} t_{k-1,2l}}, \quad k = 2, \ldots, q, \quad l = 1, \ldots, 2^{q+1-k}, \\
t_{1l} & = & t, \quad l = 1, \ldots, b, \\
t_{1l} & = & x, \quad l = b+1, \ldots, 2^q - a + b, \\
t_{1l} & = & 1, \quad l = 2^q - a + b + 1, \ldots, 2^q.
\end{array}$$

Proof: The proof is similar to that of Lemma 2 but more tedious. See [5] page 105-107 for details of the construction. Note that we defined q sets of new (dummy) variables and assigned specific values to the level 1 variables, t_{1l}, $l = 1, \ldots, 2^q$. Note also that the above construction is not unique; there are many ways to combine variables at each level. When $a = b$, the constraint $\vartheta x^{\frac{a}{b}} = \vartheta x \le t$ is linear. ■

2.4 SOCP reformulation of the active portfolio selection problem

The robust active portfolio selection problem is given by

$$\begin{array}{rcll}
v_r^* & \equiv & \text{maximize} & v_r(\boldsymbol{\phi}) = \min_{\{\boldsymbol{\alpha} \in S_\alpha, \mathbf{V}_\alpha \in S_v, \mathbf{D}_\alpha \in S_d\}} \left\{ \frac{\boldsymbol{\alpha}'\boldsymbol{\phi}}{\sqrt{\boldsymbol{\phi}'(\mathbf{V}'\mathbf{F}\mathbf{V}+\mathbf{D})\boldsymbol{\phi}}} \right\}, \\
& & \text{subject to} & \boldsymbol{\phi} \in \Phi,
\end{array} \tag{26}$$

The results in the previous section imply that for transaction cost functions of the form (25) the constraint set Φ can be expressed as a collection of SOC constraints. In this section, we use results in [15] to show that when $v_r^* > 0$ the optimization problem (26) can be reformulated as an SOCP.

Define homogeneous extension $\tilde{\Phi}$ of Φ as follows:

$$\begin{array}{rcl}
\tilde{\Phi} & \equiv & \mathbf{cl}\Big(\{ (\boldsymbol{\phi}, \zeta) : \zeta > 0, \frac{1}{\zeta}\boldsymbol{\phi} \in \Phi \} \Big), \\
& = & \mathbf{cl}\left(\left\{ (\boldsymbol{\phi}, \gamma) : \begin{array}{l}
w - \mathbf{1}'\boldsymbol{\phi} = 0, \quad w - \boldsymbol{\beta}'\boldsymbol{\phi} = 0 \quad \boldsymbol{\phi} + \mathbf{y} - \mathbf{z} = \zeta\bar{\boldsymbol{\phi}}, \\
\mathbf{1}'\boldsymbol{\phi} + \zeta T(\zeta^{-1}\mathbf{z}, \zeta^{-1}\mathbf{y}) \le \zeta \mathbf{1}'\bar{\boldsymbol{\phi}}, \\
-w\mathbf{v} \le \boldsymbol{\phi} - w\mathbf{u}, \quad \mathbf{1}'(\mathbf{y} - \mathbf{z}) \le \theta w, \\
\mathbf{z}, \ \mathbf{y}, \ w, \ \zeta \ > 0.
\end{array} \right\} \right).
\end{array} \tag{27}$$

Note that $\zeta T(\zeta^{-1}\mathbf{z}, \zeta^{-1}\mathbf{y})$ is a convex function of $(\mathbf{z}, \mathbf{y}, \zeta)$ for $\zeta > 0$ whenever T is a convex function. Moreover, for transaction cost functions of the form discussed in the previous section the constraint $\zeta T(\zeta^{-1}\mathbf{z}, \zeta^{-1}\mathbf{y}) \le \tau$ reduces to a collection of SOC constraints. For example, when T is given by the simple two-term cost function (21),

$$\tilde{\Phi} = \left\{ (\boldsymbol{\phi}, \zeta) : \begin{array}{l} w - \mathbf{1}'\boldsymbol{\phi} = 0, \quad w - \boldsymbol{\beta}'\boldsymbol{\phi} = 0 \quad \boldsymbol{\phi} + \mathbf{y} - \mathbf{z} = \zeta\bar{\boldsymbol{\phi}}, \quad \mathbf{1}'(\mathbf{y} - \mathbf{z}) \leq \theta w, \\ \mathbf{1}'\boldsymbol{\phi} + \sum_{i=1}^{n} \tau_i \leq \zeta \mathbf{1}'\bar{\boldsymbol{\phi}}, \\ \theta_{i1}(z_i + y_i) \leq \tau_i, \quad i = 1, \ldots, n \\ \left\| \begin{matrix} 2\sqrt{\vartheta_i}(z_i + y_i) \\ \tau_i - \kappa_i \end{matrix} \right\| \leq \tau_i + \kappa_i, \quad i = 1, \ldots, n, \\ \left\| \begin{matrix} 2\kappa_i \\ z_i + y_i - \zeta \end{matrix} \right\| \leq z_i + y_i + \zeta, \quad i = 1, \ldots, n, \\ -w\mathbf{v} \leq \boldsymbol{\phi} \leq w\mathbf{u}, \qquad \mathbf{z},\ \mathbf{y},\ \boldsymbol{\kappa},\ \boldsymbol{\tau},\ w,\ \zeta\ \geq 0. \end{array} \right\} \tag{28}$$

Lemma 4 *Let*

$$\begin{array}{rcll} \tilde{v}_r^* & \equiv & \textit{minimize} & \max_{\mathbf{V}_\alpha \in S_v, \mathbf{D}_\alpha \in S_d\}} \{\boldsymbol{\phi}'(\mathbf{V}'\mathbf{F}\mathbf{V} + \mathbf{D})\boldsymbol{\phi}\}, \\ & & \textit{subject to} & \min_{\boldsymbol{\alpha} \in S_\alpha} \{\boldsymbol{\alpha}'\boldsymbol{\phi}\} \geq 1, \\ & & & (\boldsymbol{\phi}, \zeta) \in \tilde{\Phi}. \end{array} \tag{29}$$

Then we have following possibilities.

(a) (29) is infeasible: $v_r^* \leq 0$.

(b) (29) is feasible: Let $(\tilde{\boldsymbol{\phi}}, \tilde{\zeta})$ *denote any optimal solution of (29). Then* $\tilde{\zeta} > 0$, $v_r^* > 0$ *and* $\tilde{\zeta}^{-1}\tilde{\boldsymbol{\phi}}$ *is optimal for (26).*

Proof: Suppose $v_r^* > 0$. Then there exists $\widehat{\boldsymbol{\phi}}$ such that $\min_{\boldsymbol{\alpha} \in S_\alpha} \boldsymbol{\alpha}'\widehat{\boldsymbol{\phi}} = \widehat{\zeta} > 0$ and $(\widehat{\zeta}^{-1}\widehat{\boldsymbol{\phi}}, \widehat{\zeta})$ is feasible for (29). This establishes (a).

Suppose $(\boldsymbol{\phi}, \zeta)$ is feasible for (29) and $\zeta = 0$. Then $\min_{\boldsymbol{\alpha} \in S_\alpha} \boldsymbol{\alpha}'\boldsymbol{\phi} \geq 1$ implies that $\boldsymbol{\phi} \neq 0$. Assume, for concreteness, that $\tilde{\Phi}$ is given by (28). The proof will proceed in an similar manner for general $\tilde{\Phi}$. From the constraints, it follows that

$$0 \leq w = \mathbf{1}'\boldsymbol{\phi} = -\sum_{i=1}^{n} \tau_i \leq 0.$$

Thus, $w = 0$ and $\tau_i = 0$, for all $i = 1, \ldots, n$. Therefore, $\mathbf{z} = \mathbf{y} = 0$; hence, $\boldsymbol{\phi} = \mathbf{y} - \mathbf{z} = \mathbf{0}$. A contradition. Thus, we must have $\zeta > 0$ for all $(\boldsymbol{\phi}, \zeta)$ feasible for (29).

Suppose $(\tilde{\boldsymbol{\phi}}, \tilde{z})$ is optimal for (29). Then $(\tilde{\zeta})^{-1}\tilde{\boldsymbol{\phi}} \in \Phi$. Therefore,

$$\begin{aligned} v_r^* &\geq \min_{\{\boldsymbol{\alpha} \in S_\alpha, \mathbf{V}_\alpha \in S_v, \mathbf{D}_\alpha \in S_d\}} \left\{ \frac{\boldsymbol{\alpha}'(\tilde{\zeta}^{-1}\tilde{\boldsymbol{\phi}})}{\sqrt{(\tilde{\zeta}^{-1}\tilde{\boldsymbol{\phi}})'(\mathbf{V}'\mathbf{F}\mathbf{V} + \mathbf{D})(\tilde{\zeta}^{-1}\tilde{\boldsymbol{\phi}})}} \right\} \\ &= \min_{\{\boldsymbol{\alpha} \in S_\alpha, \mathbf{V}_\alpha \in S_v, \mathbf{D}_\alpha \in S_d\}} \left\{ \frac{\boldsymbol{\alpha}'\tilde{\boldsymbol{\phi}}}{\sqrt{\tilde{\boldsymbol{\phi}}'(\mathbf{V}'\mathbf{F}\mathbf{V} + \mathbf{D})\tilde{\boldsymbol{\phi}}}} \right\} = \frac{1}{\sqrt{\tilde{v}_r^*}} > 0, \end{aligned}$$

where the first equality follows from the fact that the objective $v_r(\boldsymbol{\phi})$ is homogeneous with degree-0.

Suppose $v_r^* > 1/\sqrt{\tilde{v}_r^*}$. Let $\widehat{\boldsymbol{\phi}} \in \Phi$ denote any optimal solution of (26). Since $v_r^* > 0$, it follows that $\widehat{\zeta} = \min_{\boldsymbol{\alpha} \in S_\alpha} \boldsymbol{\alpha}'\widehat{\boldsymbol{\phi}} > 0$. It is easy to check that $(\widehat{\zeta}^{-1}\widehat{\boldsymbol{\phi}}, \widehat{\zeta})$ is feasible for (29) and

$$\max_{\mathbf{V}_\alpha \in S_v, \mathbf{D}_\alpha \in S_d\}} \{(\widehat{\zeta}^{-1}\widehat{\boldsymbol{\phi}})'(\mathbf{V}'\mathbf{F}\mathbf{V} + \mathbf{D})(\widehat{\zeta}^{-1}\widehat{\boldsymbol{\phi}}\} = \frac{1}{(v_r^*)^2} < \tilde{v}_r^*.$$

A contradiction. Therefore, it follows that $v_r^* = \frac{1}{\sqrt{\tilde{v}_r^*}}$ and $\tilde{\zeta}^{-1}\tilde{\boldsymbol{\phi}}$ is optimal for (26). This establishes (b). ■

In practice, we rebalance only if there exists a new portfolio $\boldsymbol{\phi} \in \Phi$ with $v_r(\boldsymbol{\phi}) > 0$. Therefore, for all practical purposes, the two optimization problems are equivalent. It is clear that (29) is equivalent to

$$\begin{array}{lllll} \text{minimize} & \nu + \gamma & & & \\ \text{subject to} & \boldsymbol{\phi}\mathbf{V}'\mathbf{F}\mathbf{V}\boldsymbol{\phi} & \leq & \nu, & \text{for all } \mathbf{V} \in S_v, \\ & \boldsymbol{\phi}'\mathbf{D}\boldsymbol{\phi} & \leq & \gamma, & \text{for all } \mathbf{D} \in S_d, \\ & \boldsymbol{\alpha}'\boldsymbol{\phi} & \geq & 1, & \text{for all } \boldsymbol{\alpha} \in S_\alpha, \\ & (\boldsymbol{\phi}, \zeta) & \in & \tilde{\Phi}. & \end{array} \tag{30}$$

From the definition of S_α and S_d it follows that

$$\boldsymbol{\alpha}'\boldsymbol{\phi} \geq 1, \; \forall\; \boldsymbol{\alpha} \in S_\alpha \quad \Leftrightarrow \quad \boldsymbol{\alpha}_0'\boldsymbol{\phi} - \boldsymbol{\eta}'\,|\boldsymbol{\phi}| \geq 1, \quad \Leftrightarrow \quad \boldsymbol{\alpha}_0'\boldsymbol{\phi} - \boldsymbol{\eta}'\boldsymbol{\psi}, \boldsymbol{\psi} \geq \boldsymbol{\phi}, \boldsymbol{\psi} \geq -\boldsymbol{\phi}, \tag{31}$$

where $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_n)'$, and

$$\boldsymbol{\phi}'\mathbf{D}\boldsymbol{\phi} \leq \gamma, \; \forall\; \mathbf{D} \in S_d \quad \Leftrightarrow \quad \boldsymbol{\phi}'\,\mathbf{diag}(\bar{\mathbf{d}} + \boldsymbol{\delta})\boldsymbol{\phi} \leq \gamma \quad \Leftrightarrow \quad \left\| \begin{bmatrix} 2\,\mathbf{diag}(\bar{\mathbf{d}} + \boldsymbol{\delta})^{1/2}\boldsymbol{\phi} \\ 1 - \gamma \end{bmatrix} \right\| \leq 1 + \gamma, \tag{32}$$

where $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_n)'$. In [15], it was shown that for fixed portfolio $\widehat{\boldsymbol{\phi}}$ and a number ν, $\widehat{\boldsymbol{\phi}}'\mathbf{V}'\mathbf{F}\mathbf{V}\widehat{\boldsymbol{\phi}} \leq \nu$, for all $\mathbf{V} \in S_v$ if, and only if, there exist $\sigma, \varsigma \geq 0$ and $\mathbf{h} \in \mathbf{R}_+^m$ satisfying

$$\begin{array}{rcll} \varsigma + \mathbf{1}'\mathbf{h} & \leq & \nu, & \\ \sigma & \leq & \frac{1}{\lambda_{\max}(\mathbf{H})}, & \\ \left\| \begin{bmatrix} 2r \\ \sigma - \varsigma \end{bmatrix} \right\| & \leq & \sigma + \varsigma, & \\ \left\| \begin{bmatrix} 2g_i \\ 1 - \sigma\lambda_i - h_i \end{bmatrix} \right\| & \leq & 1 - \sigma\lambda_i + h_i, & i = 1, \ldots, m, \end{array} \tag{33}$$

where $r = \boldsymbol{\rho}'|\boldsymbol{\phi}|$, $\mathbf{Q}\Lambda\mathbf{Q}'$ is the spectral decomposition of $\mathbf{H} = \mathbf{G}^{-\frac{1}{2}}\mathbf{F}\mathbf{G}^{-\frac{1}{2}}$, $\Lambda = \mathbf{diag}(\boldsymbol{\lambda})$, and $\mathbf{g} = \mathbf{Q}'\mathbf{H}^{\frac{1}{2}}\mathbf{G}^{\frac{1}{2}}\mathbf{V}_0\boldsymbol{\phi}$. From (31), (32) and (33), it follows that the robust maximum information ratio problem (26) is equivalent to an SOCP. Since the non-robust active portfolio selection problem is a special case of the robust problem where the uncertainty sets are all singletons, it follows the non-robust problem is also equivalent to an SOCP.

The fact that the active portfolio selection problems are SOCPs has important theoretical and practical implications. Since the computational complexity of an SOCP is comparable to that of a convex quadratic program, it follows that robust active portfolio selection is able to provide protection against parameter fluctuations at very moderate computational cost. Morever, a number commercial solvers such as MOSEK, CPLEX and Frontline System (supplier of EXCEL SOLVER) provide the capability for solving SOCPs in a numerically robust manner.

3 Alternative models for robust active portfolio management

3.1 Side constraints on alpha

In §2.1, we assumed that the random exceptional return $\Delta\mathbf{r}$ is given by the robust factor model

$$\Delta\mathbf{r} = \boldsymbol{\alpha} + \mathbf{V}'\mathbf{f} + \boldsymbol{\epsilon}$$

where $\boldsymbol{\alpha} \in S_\alpha$, $\mathbf{V} \in S_v$, $\mathbf{f} \sim \mathcal{N}(\mathbf{0}, \mathbf{F})$ and $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{D})$, and the sets S_α, S_v and S_d are defined in (6), (5) and (4), respectively. In §2.4 we further assumed that the uncertain parameters $(\boldsymbol{\alpha}, \mathbf{V}, \mathbf{D})$ could be *independently* chosen to minimize the value of the information ratio $\mathrm{IR}(\boldsymbol{\phi})$. This assumption leads to very pessimistic estimate of the performance of the portfolio $\boldsymbol{\phi}$, and as a consequence (29) can become infeasible even when there exist portfolios with reasonable active returns. This pessimistic behavior can be controlled by suitably constraining the set of feasible market parameters. In this section, we survey some of the constraints that still allow the portfolio selection problem to be recast as an SOCP.

3.1.1 Net zero alpha

A net zero alpha constraint requires that the return vector $\boldsymbol{\alpha}$ satisfy $\sum_{i\in I} \alpha_i = \sum_{i \in I} \alpha_{0i}$ for some set of indices $I \subseteq \{1, \ldots, n\}$ (e.g. stocks in a particular industry). Such a constraint ensures that, although an individual α_i, $i \in I$, may differ from its nominal value α_{0i}, the average deviation is equal to zero. Moreover, the robust constraint $\min_{\boldsymbol{\alpha}\in S_\alpha} \boldsymbol{\alpha}'\boldsymbol{\phi} \geq 1$ now becomes

$$\boldsymbol{\alpha}_0'\boldsymbol{\phi} + \min\left\{\boldsymbol{\xi}'\boldsymbol{\phi} : \sum_{i\in I}\xi_i = 0, |\boldsymbol{\xi}| \leq \boldsymbol{\eta}\right\} = \boldsymbol{\alpha}_0'\boldsymbol{\phi} - \sum_{i\notin I}\eta_i\,|\phi_i| - \sum_{i\in I}\eta_i\,|\phi_i - \lambda| \geq 1.$$

where λ is a new decision variable. Thus, net zero alpha constraints are modeled by linear constraints.

3.1.2 Incorporating analyst views

We propose two methods for incorporating analyst views into robust active portfolio selection. The first relies on Bayesian analysis. The bounds $\boldsymbol{\eta}$ defining the uncertainty set S_α were set using the confidence level ω and the posterior density

$$\boldsymbol{\alpha} = \mathcal{N}(\boldsymbol{\alpha}_0, \boldsymbol{\Sigma}), \tag{34}$$

implied by historical data. Suppose, in addition, analysts provide the portfolio manager with private views about the future $\boldsymbol{\alpha}$. For example, the analysts may believe that $\mathbb{P}(\alpha_i - \alpha_j \geq \pi) = \eta$. This view can be expressed by setting

$$\alpha_i - \alpha_j = \kappa + \sigma\zeta,$$

where $\zeta \sim \mathcal{N}(0, 1)$, the parameters (κ, σ) are chosen to satisfy $\mathcal{F}_{\mathcal{N}}(\frac{\kappa-\pi}{\sigma}) = \eta$ and $\mathcal{F}_{\mathcal{N}}$ denotes the CDF of a standard Normal random variable. (Note that there is some flexibility in the choice of (κ, σ).) Thus, if the portfolio manager receives m different analyst views of the form $\mathbb{P}(\mathbf{p}'\alpha \geq \pi) = \eta$, they can collectively be represented as

$$\mathbf{P}\boldsymbol{\alpha} = \mathcal{N}(\boldsymbol{\kappa}, \boldsymbol{\Lambda}). \tag{35}$$

for suitably chosen $\mathbf{P} \in \mathbb{R}^{m\times n}$, $\boldsymbol{\kappa} \in \mathbb{R}^m$ and $\boldsymbol{\Lambda} \in \mathbb{R}^{m\times m}$. From (34) and (35) it follows that the posterior density of $\boldsymbol{\alpha}$ given *both* historical data and analysts' views is given by

$$\boldsymbol{\alpha} = \mathcal{N}(\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{\Sigma}}), \tag{36}$$

where $\bar{\boldsymbol{\Sigma}} = (\boldsymbol{\Sigma}^{-1} + \boldsymbol{\Lambda}^{-1})^{-1}$ and $\bar{\boldsymbol{\alpha}} = \bar{\boldsymbol{\Sigma}}(\boldsymbol{\Sigma}^{-1}\boldsymbol{\alpha}_0 + \mathbf{P}'\boldsymbol{\Lambda}^{-1}\boldsymbol{\kappa})$. From (36), it follows that the ω-confidence region for $\boldsymbol{\alpha}$ is given by s'

$$S_\alpha = \{\alpha : (\boldsymbol{\alpha} - \bar{\boldsymbol{\alpha}})'\bar{\boldsymbol{\Sigma}}^{-1}(\boldsymbol{\alpha} - \bar{\boldsymbol{\alpha}}) \leq \boldsymbol{\chi}_n^{-1}(\omega)\},$$

where $\boldsymbol{\chi}_n$ denotes the CDF of a χ^2 random variable with n degrees of freedom. Thus, the robust constraint $\min_{\boldsymbol{\alpha}\in S_\alpha}\{\boldsymbol{\alpha}'\boldsymbol{\phi}\} \geq 1$ is equivalent to the SOC-constraint

$$\bar{\boldsymbol{\alpha}}'\boldsymbol{\phi} - \boldsymbol{\chi}_n^{-1}(\omega)\|\bar{\boldsymbol{\Sigma}}^{\frac{1}{2}}\boldsymbol{\phi}\| \geq 1.$$

This method of incorporating views is identical to that used in the Black-Litterman model [7] – with the added feature of using second-order information to set confidence levels.

In practice, however, it may be difficult to elicit the vector $\boldsymbol{\kappa}$ and the covariance $\boldsymbol{\Lambda}$. Our second method assumes that analysts' views are of the form

$$l \leq \mathbf{p}'\boldsymbol{\alpha} \leq u$$

or equivalently

$$\mathbf{p}'\boldsymbol{\alpha} = \frac{u+l}{2} + v, \quad |v| \leq \frac{u-l}{2}.$$

Suppose the manager receives m analysts' views. These can be expressed as $|\mathbf{P}\boldsymbol{\alpha} - \boldsymbol{\kappa}| \leq \boldsymbol{\lambda}$, where $\mathbf{P} \in \mathbb{R}^{m\times n}$ and $\boldsymbol{\kappa}, \boldsymbol{\lambda} \in \mathbb{R}^m$. Coupling this constraint with the bounds $|\boldsymbol{\alpha} - \boldsymbol{\alpha}_0| \leq \boldsymbol{\eta}$ implied by historical data results in the uncertainty set

$$\begin{aligned} S_\alpha &= \{\boldsymbol{\alpha} : |\boldsymbol{\alpha} - \boldsymbol{\alpha}_0| \leq \boldsymbol{\eta}, |\mathbf{P}\boldsymbol{\alpha} - \boldsymbol{\kappa}| \leq \boldsymbol{\lambda}\}, \\ &= \{\boldsymbol{\alpha} : \boldsymbol{\alpha} = \boldsymbol{\alpha}_0 + \mathbf{u}, \mathbf{P}\mathbf{u} - \mathbf{v} = \boldsymbol{\kappa} - \mathbf{P}\boldsymbol{\alpha}_0, |\mathbf{u}| \leq \boldsymbol{\eta}, |\mathbf{v}| \leq \boldsymbol{\lambda}\}. \end{aligned} \tag{37}$$

This uncertainty set is completely defined by setting the center $\boldsymbol{\kappa}$ and the bounds $\boldsymbol{\lambda}$. We believe that eliciting these two parameters would be significantly easier than eliciting the covariance matrix $\boldsymbol{\Lambda}$.

From (37), it follows that the robust constraint $\min_{\boldsymbol{\alpha}\in S_\alpha}\{\boldsymbol{\alpha}'\boldsymbol{\phi}\} \geq 1$ is equivalent to

$$\boldsymbol{\alpha}_0'\boldsymbol{\phi} + (\mathbf{P}\boldsymbol{\alpha}_0 - \boldsymbol{\kappa})'\boldsymbol{\gamma} - \boldsymbol{\eta}' \left|\boldsymbol{\phi} - \mathbf{P}'\boldsymbol{\gamma}\right| - \boldsymbol{\lambda}' |\boldsymbol{\gamma}| \geq 1,$$

where $\boldsymbol{\gamma}$ is a new decision variable. It is easy to see that the above constraint can be *linearized* and, therefore, can be incorporated into an SOCP.

3.2 Data driven methods

All the methods discussed in §2.2 and §3.1 are *parametric* methods in the sense that return data is assumed to be distributed according a known distribution but with uncertain parameters. In this section, we describe data-driven robust methods where we do not make any parametric assumptions about the return distribution. In the first subsection below, we consider non-parametric shortfall minimization. In the subsequent subsections, we introduce robustness into the shortfall minimization problem and discuss tractability issues.

3.2.1 Non-parametric shortfall minimization

Let $\{\mathbf{D}^{(k)} : k = 1, \ldots, N\}$ and $\{B^{(k)} : k = 1, \ldots, N\}$ denote, respectively, the historical data on the return $\mathbf{r} \in \mathbb{R}^n$ of the n assets in the market and the return of the benchmark. Recall that in §2.2 we argued that a natural objective for an active portfolio manager is to minimize the shortfall probability $\mathbb{P}(\mathbf{r}'\boldsymbol{\phi} \leq wr_b)$ (see, e.g. (14)). Since we only have the raw historical data, we approximate the true distribution $\mathbb{P}$ by the empirical measure $\widehat{\mathbb{P}}$. We have N samples

$$\gamma_k(\boldsymbol{\phi}) = wB^{(k)} - (\mathbf{D}^{(k)})'\boldsymbol{\phi}, \quad k = 1, \ldots, N, \tag{38}$$

for the random shortfall of the portfolio $\boldsymbol{\phi}$. Therefore, the active manager's optimization problem reduces to

$$\begin{array}{ll} \text{minimize} & \frac{1}{N}\sum_{k=1}^N \boldsymbol{\chi}\big(\gamma_k(\boldsymbol{\phi})\big), \\ \text{subject to} & \boldsymbol{\phi} \in \Phi, \end{array} \tag{39}$$

where $\boldsymbol{\chi}(\cdot)$ denotes the indicator function of the positive axis, i.e. $\boldsymbol{\chi}(x) = 1$ for $x \geq 0$ and zero, otherwise. The optimization problem (39) is equivalent to a mixed 0-1 SOCP with one knapsack constraint, and hence, is not a convex problem.

The *value-at-risk* ($\mathrm{VaR}_\epsilon(\boldsymbol{\phi})$) of a portfolio $\boldsymbol{\phi}$ at probability ϵ is defined as

$$\mathrm{VaR}_\epsilon(\boldsymbol{\phi}) = \inf_x \left\{ \mathbb{P}((r_b\mathbf{1} - \mathbf{r})'\boldsymbol{\phi} \geq x) \leq \epsilon \right\}.$$

It follows that $\mathbb{P}((r_b\mathbf{1} - \mathbf{r})'\boldsymbol{\phi} \leq 0) \leq \epsilon$ if, and only if, $\mathrm{VaR}_\epsilon(\boldsymbol{\phi}) \leq 0$. Thus, (39) can be interpreted as maximizing the probability ϵ for which $\mathrm{VaR}_\epsilon(\boldsymbol{\phi}) \leq 0$.

We next show how to construct a tractable convex approximation for (39). Since $\boldsymbol{\chi}(x) \leq (vx+1)^+$ for all $v > 0$, it follows that

$$\frac{1}{N}\sum_{k=1}^{N} \boldsymbol{\chi}\big(\gamma_k(\boldsymbol{\phi})\big) \leq \frac{1}{N}\sum_{k=1}^{N} \Big(\gamma_k(v\boldsymbol{\phi}) + 1\Big)^+ \tag{40}$$

Let $\boldsymbol{\psi} = v\boldsymbol{\phi}$. Then $(\boldsymbol{\psi}, v) \in \tilde{\Phi}$ (see (27)). From (40) it follows that

$$\begin{array}{rl} \text{minimize} & \frac{1}{N}\sum_{i=1}^{N}\big(\gamma_k(\boldsymbol{\psi}) + 1\big)^+, \\ \text{subject to} & (\boldsymbol{\psi}, v) \in \tilde{\Phi}, \end{array} \tag{41}$$

is a convex approximation of the active manager's problem. Since $\mathbb{P}((r_b\mathbf{1} - \mathbf{r})'\boldsymbol{\phi} \geq 0) \leq \mathbb{E}(v(r_b\mathbf{1} - \mathbf{r})'\boldsymbol{\phi} + 1)^+$, it follows that

$$\mathbb{E}((r_b\mathbf{1} - \mathbf{r})'\boldsymbol{\phi} + v^{-1})^+ - v^{-1}\epsilon \leq 0 \tag{42}$$

implies that $\mathbb{P}((r_b\mathbf{1} - \mathbf{r})'\boldsymbol{\phi} \geq 0) \leq \epsilon$, or equivalently, $\mathrm{VaR}_\epsilon(\boldsymbol{\phi}) \leq 0$. Since the left hand side of the expression in (42) is positive for $t < 0$, it follows that

$$\inf_t \left\{ t + \frac{1}{\epsilon}\mathbb{E}((r_b\mathbf{1} - \mathbf{r})'\boldsymbol{\phi} - t)^+ \right\} \leq 0,$$

implies that $\mathrm{VaR}_\epsilon(\boldsymbol{\phi}) \leq 0$. Rockafellar and Uryasev [34] have shown that

$$\mathrm{CVaR}_\epsilon(\boldsymbol{\phi}) = \mathbb{E}[(r_b\mathbf{1} - \mathbf{r})'\boldsymbol{\phi} \mid (r_b\mathbf{1} - \mathbf{r})'\boldsymbol{\phi} \geq \mathrm{VaR}_\epsilon(\boldsymbol{\phi})] = \inf_t \left\{ t + \frac{1}{\epsilon}\mathbb{E}((r_b\mathbf{1} - \mathbf{r})'\boldsymbol{\phi} - t)^+ \right\}. \tag{43}$$

From (43) and (42) it follows that $\mathrm{CVaR}_\epsilon(\boldsymbol{\phi})$ is a conservative convex approximation for $\mathrm{VaR}_\epsilon(\boldsymbol{\phi})$ and that (40) can be interpreted as maximizing the probability ϵ for which $\mathrm{CVaR}_\epsilon(\boldsymbol{\phi}) \leq 0$. Nemirovski and Shapiro [30] show that the CVaR is, in fact, the *tightest* convex approximation of the VaR. Unlike VaR, CVaR is a *coherent* risk measure, and is becoming popular in risk management applications. Thus, the sample approximation (40) is an important approximation problem in its own right.

Let $\boldsymbol{\phi}^*$ denote the portfolio that minimizes the true CVaR (note that $\boldsymbol{\phi}^*$ is not computable in any practical sense). Let $\widehat{\boldsymbol{\phi}}$ denote the optimal solution of the sample-based CVaR optimization problem (40). Then two important question emerge:

(i) How close is the sample approximation of the CVaR to the true CVaR of the portfolio $\widehat{\boldsymbol{\phi}}$?

(ii) How close is the CVaR of $\widehat{\boldsymbol{\phi}}$ to the CVaR of the true optimal $\boldsymbol{\phi}^*$?

Both of these questions can be answered using results from statistical learning theory and the Vapnik-Chervonenkis dimension. See [20] for details.

A loss function ρ that maps real valued random variables X to real numbers is called a *Chocquet* loss function if there exists a non-decreasing convex function $v : [0,1] \mapsto [0,1]$ such that

1. $v(0) = 0$ and $v(1) = 1$
2. $\rho(X) = \int_0^1 F_X^{-1}(t)dv(t)$, where F_X denotes the CDF of the random variable X.

Since the

$$\mathrm{CVaR}_\epsilon(X) = \mathbb{E}\big[X : X \geq \mathrm{VaR}_\epsilon(X)\big] = \int_{1-\epsilon}^1 F_X^{-1}(t)\frac{dt}{\epsilon} = \int_0^1 F_X^{-1}(t)d\widehat{v}(t)$$

for $\widehat{v}(t) = \frac{1}{\epsilon}(t-1+\epsilon)^+$, it follows that CVaR_ϵ is a Choquet loss function.

Let $\Gamma_\phi = (r_b\mathbf{1} - \mathbf{r})'\boldsymbol{\phi}$ denote the random return of the portfolio $\boldsymbol{\phi}$ below the benchmark r_b and let ρ be any Choquet loss function. Then another natural optimization problem for the active portfolio manager is given by

$$\begin{array}{rl} \text{minimize} & \rho(\gamma_\phi), \\ \text{subject to} & \boldsymbol{\phi} \in \Phi. \end{array} \tag{44}$$

The sample approximation (40) for the CVar optimization problem can be extended to all Choquet loss functions as follows.

$$\begin{aligned} \rho(\Gamma_\phi) &= \int_0^1 F_{\Gamma_\phi}^{-1}(t)dv(t), \\ &\approx \int_0^1 \widehat{F}_{\Gamma_\phi}^{-1}(t)dv(t), \\ &= \sum_{k=1}^N q_k f_{[k]}(\phi) \end{aligned}$$

where $\widehat{F}_{\Gamma_\phi}$ denotes the empirical CDF of Γ_ϕ constructed from the N samples $\boldsymbol{\gamma}(\boldsymbol{\phi}) = (\gamma_1(\boldsymbol{\phi}), \ldots, \gamma_N(\boldsymbol{\phi}))'$ of the shortfall (see (38)),

$$q_k = \begin{cases} v\big((N-k+1)/N\big) - v\big((N-k-1)/N\big), & k = 1, 2, \ldots, N-1, \\ v\big(1/N\big), & \text{otherwise,} \end{cases}$$

and $f_{[k]}(\boldsymbol{\phi})$ denotes the sum of the k largest terms in $\boldsymbol{\gamma}(\boldsymbol{\phi})$. Thus, (44) can be approximated by the following sample-based optimization

$$\begin{array}{rl} \text{minimize} & \sum_{k=1}^N q_k f_{[k]}(\boldsymbol{\phi}), \\ \text{subject to} & \boldsymbol{\phi} \in \Phi. \end{array} \tag{45}$$

Berstimas and Brown [6] use linear programming duality to show that

$$f_{[k]}(\boldsymbol{\phi}) \quad = \quad \begin{array}{rl} \text{maximize} & \mathbf{a}'\boldsymbol{\gamma}(\boldsymbol{\phi}), \\ \text{subject to} & \mathbf{1}'\mathbf{a} = k, \\ & \mathbf{0} \leq \mathbf{a} \leq \mathbf{1}, \end{array} \quad = \quad \begin{array}{rl} \text{minimize} & k\tau^{(k)} + \mathbf{1}'\mathbf{u}^{(k)}, \\ \text{subject to} & \tau^{(k)}\mathbf{1} + \mathbf{u}^{(k)} \geq \boldsymbol{\gamma}(\boldsymbol{\phi}), \\ & \mathbf{u}^{(k)} \geq \mathbf{0}, \end{array} \tag{46}$$

Since $q \geq 0$, the dual formulation in (46) implies that (45) is equivalent to the SOCP

$$\begin{array}{rl} \text{minimize} & \sum_{k=1}^{N} q_k(k\tau^{(k)} - \mathbf{1}'\mathbf{u}^{(k)}), \\ \text{subject to} & \tau^{(k)}\mathbf{1} - \mathbf{u}^{(k)} - \boldsymbol{\gamma}(\boldsymbol{\phi}) \geq \mathbf{0}, \\ & \mathbf{u}^{(k)} \geq \mathbf{0}, \quad k = 1, \ldots, N, \\ & \boldsymbol{\phi} \in \Phi. \end{array} \tag{47}$$

3.2.2 Choquet loss functions with parameter uncertainty

In this section, we use results in §2.1 to extend the above data-driven model (47) to include parameter uncertainty.

From (12), it follows that

$$(r_b\mathbf{1} - \mathbf{r})'\boldsymbol{\phi} = \Delta\mathbf{r}'\boldsymbol{\phi} = (\boldsymbol{\alpha} + \mathbf{V}'\mathbf{f})'\boldsymbol{\phi}$$

where $\mathbf{f}$ denotes the random return on the factors driving the market. Suppose $\boldsymbol{\alpha}$ and $\mathbf{V}$ are uncertain with uncertainty sets S_α and S_v respectively, then the k-th sample $\gamma_k(\boldsymbol{\phi}) = (B^{(k)}\mathbf{1} - \mathbf{D}^{(k)})'\boldsymbol{\phi}$ of the shortfall of the portfolio $\boldsymbol{\phi}$ is uncertain with the uncertainty

$$\Delta\gamma_k = -\Delta\boldsymbol{\alpha}'\boldsymbol{\phi} + (\mathbf{f}^{(k)})'\Delta\mathbf{V}\boldsymbol{\phi},$$

where $\mathbf{f}^{(k)}$ denotes the k-th sample of the factor returns. We interpret the uncertainty $\Delta\gamma_k$ as follows: $\gamma_k(\boldsymbol{\phi})$ is a sample of the shortfall in the past; in the future, however, the value of the parameters $\boldsymbol{\alpha}$ and $\mathbf{V}$ may shift to some other value in their associated uncertainty sets and the "correct" sample of the new shortfall distribution is given by $\gamma_k(\boldsymbol{\phi}) + \Delta\gamma_k$. Therefore, in order to protect against all possible parameter shifts one should solve the following robust version of (47)

$$\begin{array}{rl} \text{minimize} & \sum_{k=1}^{N} q_k(k\tau^{(k)} - \mathbf{1}'\mathbf{u}^{(k)}), \\ \text{subject to} & \tau^{(k)} - \mathbf{u}_j^{(k)} - \gamma_j(\boldsymbol{\phi}) - \beta - \nu_j \geq 0, \quad j, k = 1, \ldots, N, \\ & \max_{\boldsymbol{\alpha}_0 + \Delta\boldsymbol{\alpha} \in S_\alpha}\{-\Delta\boldsymbol{\alpha}'\boldsymbol{\phi}\} \leq \beta, \\ & \max_{\mathbf{V}_0 + \Delta\mathbf{V} \in S_v}\{(\mathbf{f}^{(j)})'\Delta\mathbf{V}\boldsymbol{\phi}\} \leq \nu_j, \quad j = 1, \ldots, N, \\ & \mathbf{u}^{(k)} \geq \mathbf{0}, \quad k = 1, \ldots, N, \\ & \boldsymbol{\phi} \in \Phi. \end{array} \tag{48}$$

Both (47) and (48) suffer from the drawback that the number of variables in the problem increasse as $O(N^2)$, where N is the number of samples. Next, we describe a more tractable formulation for a restricted class of Choquet loss functions.

Let $v_p(t) = \frac{1}{p}(t - 1 + p)^+$. Then it is easy to check that any non-decreasing convex function with $v(0) = 1 - v(1) = 0$ can be represented as

$$v(t) = \int_0^1 v_p(t) m(dp)$$

where m denotes a probability measure on $[0,1]$. Consider the restricted class of Choquet loss ρ functions such that the corresponding measure m_ρ is discrete, i.e. the corresponding convex function v_ρ is of the form

$$v_\rho(t) = \sum_{s=1}^{S} m_s v_{p_s}(t)$$

Since the loss function ρ_s corresponding to $v_{p_s}(t)$ is CVaR_{p_s}, it follows that $\rho_s(\boldsymbol{\phi}) = \inf_t\{t + \frac{1}{p_s}\mathbb{E}((r_b\mathbf{1} - \mathbf{r})'\boldsymbol{\phi} - t)^+$, and the loss function ρ corresponding to v_ρ is $\rho = \sum_{s=1}^{S} m_s \mathrm{CVaR}_{p_s}$. Thus, we have that for this restricted class of loss functions the robust problem (48) can be reformulated as follows

$$\begin{array}{rl} \text{minimize} & \sum_{s=1}^{S} m_s\beta_s, \\ \text{subject to} & t_s - \frac{1}{p_s N}\sum_{j=1}^{N}(\beta + \eta_j - t_s)^+ \leq \beta_s, \\ & \max_{\Delta\boldsymbol{\alpha}\in S_\alpha}\{-\Delta\boldsymbol{\alpha}'\boldsymbol{\phi}\} \leq \beta, \\ & \gamma_j(\boldsymbol{\phi}) + \max_{\Delta\mathbf{V}\in S_v}\{(\mathbf{f}^{(j)})'\Delta\mathbf{V}\boldsymbol{\phi}\} \leq \nu_j, \quad j = 1,\ldots,N, \\ & \boldsymbol{\phi} \in \Phi. \end{array} \tag{49}$$

This program has only $O(NS)$ variables. Thus, this formulation become much more attractive when $S = o(N)$.

3.2.3 Choquet loss function with non-parametric uncertainty

In this section, we describe robustness with respect to a non-parametric set of measures. This measure of robustness is formulated in terms of the following Stackleberg game between the investor and nature:

1. Investor choose a portfolio $\boldsymbol{\phi} \in \Phi$.
2. Nature chooses a shortfall distribution such that the corresponding CDF F satisfies:
   (a) $\sup_{x\in\mathbb{R}} |F(x) - \widehat{F}(x)| \leq \epsilon$, where $\widehat{F}$ denotes the empirical distribution implied by the vector $\boldsymbol{\gamma}(\boldsymbol{\phi})$.
   (b) $F\big(\kappa \max_k\{\gamma_k(\boldsymbol{\phi})\}\big) = 1$.

The constraint (b) on the power of Nature is required to ensure that the set of admissible measures is *tight*. It can be replaced by any other constraint that ensures tightness. The Robust CVaR optimization problem is given by

$$\begin{array}{rl} \text{maximize} & \max_{F\in\mathcal{F}(\boldsymbol{\psi})} \mathbf{E}(Z+1)^+, \\ \text{subject to} & \boldsymbol{\psi} \in \tilde{\Phi}. \end{array} \tag{50}$$

where $Z \sim F$.

Fix $\boldsymbol{\phi}$ and $\boldsymbol{\gamma} = \boldsymbol{\gamma}(\boldsymbol{\phi})$. Let $\gamma_{(k)}$ denote the k-th order statistic of the vector $\boldsymbol{\gamma}$. Then for $x \in [\gamma_{(k)}, \gamma_{(k+1)})$,

$$\mid F(x) - \widehat{F}(x) \mid = \mid F(x) - k/N \mid \leq \max \left\{ \mid F(\gamma_{(k+1)}^{-}) - k/N \mid, \mid F(\gamma_{(k)}) - k/N \mid \right\},$$

where $F(x^{-})$ denotes the limit from the left. Thus, it follows that

$$\sup_{x} |F(x) - \widehat{F}(x)| = \max_{1 \leq k \leq n} \left\{ \mid F(\gamma_{(k+1)}^{-}) - k/N \mid, \mid F(\gamma_{(k)}) - k/N \mid \right\}. \tag{51}$$

For any fixed $F \in \mathcal{F}(\boldsymbol{\psi})$, we have that

$$\mathbb{E}(Z+1)^{+} \leq \sum_{k=1}^{N} \gamma_{(k)} \big(F(\gamma_{(k)}) - F(\gamma_{(k-1)}) \big) + (1+\epsilon) \gamma_{(N)} \big(F(\gamma_{(N+1)}) - F(\gamma_{(N)}) \big), \tag{52}$$

where $\gamma_{(0)} = -\infty$ and $\gamma_{(N+1)} = \kappa \gamma_{(N)}$, and the equality holds only if F puts all the probability mass on the set $\{\gamma_{(k)} : k = 1, \ldots, N\} \cup \{\kappa \gamma_{(N)}\}$. From (51) and (52), it follows that

$$\begin{array}{rll} \displaystyle\max_{F \in \mathcal{F}(\boldsymbol{\psi})} \mathbf{E}(Z+1)^{+} = & \text{maximize} & \sum_{k=1}^{N} (\gamma_{(k)} + 1)^{+} p_k + (\kappa \gamma_{(N)} + 1)^{+} p_{N+1}, \\ & \text{subject to} & \left| \sum_{j=1}^{k} p_j - k/N \right| \leq \epsilon, \quad k = 1, \ldots, N, \\ & & \sum_{k=1}^{N+1} p_k = 1, \\ & & \mathbf{p} \geq \mathbf{0}. \end{array}$$

For $\epsilon < 1$, it is easy to check that the optimal solution of this LP is given by

$$\max_{F \in \mathcal{F}(\boldsymbol{\psi})} \mathbf{E}(Z+1)^{+} = \left(\frac{m}{N} - \epsilon \right) (\gamma_{(m)} + 1)^{+} + \frac{1}{N} \sum_{k=m+1}^{N} (\gamma_{(k)} + 1)^{+} + \epsilon (\kappa \gamma_{(N)} + 1)^{+},$$

where $m = \lceil N\epsilon \rceil$. Thus, the robust active portfolio selection problem (50) is equivalent to

$$\begin{array}{ll} \text{maximize} & \left(\frac{m}{N} - \epsilon \right) \left(\sum_{k=m}^{N} (\gamma_{(k)}(\boldsymbol{\psi}) + 1)^{+} \right) + \\ & \left(\epsilon - \frac{(m-1)}{N} \right) \left(\sum_{k=m+1}^{N} (\gamma_{(k)} + 1)^{+} \right) + \epsilon (\kappa \gamma_{(N)}(\boldsymbol{\psi}) + 1)^{+}, \\ \text{subject to} & \boldsymbol{\psi} \in \tilde{\Phi}. \end{array} \tag{53}$$

Let $\tilde{\gamma}_k(\boldsymbol{\psi}) = (\gamma_k(\boldsymbol{\psi}) + 1)^{+}$. Then $\sum_{k=m}^{N} \gamma_{(k)}(\boldsymbol{\psi}) + 1)^{+} = \sum_{k=m}^{N} \tilde{\gamma}_{(k)}$ is the sum of the $N - m + 1$ largest terms in the vector $\tilde{\boldsymbol{\gamma}}(\boldsymbol{\psi})$. Thus, the dual formulation for the sum of k largest terms (see (46)) it follows that (53) is equivalent to

$$\begin{array}{ll} \text{maximize} & \left(\frac{m}{N} - \epsilon \right) \left((N - m + 1) \tau^{(1)} + \mathbf{1}' \mathbf{u}^{(1)} \right) + \\ & \left(\epsilon - \frac{(m-1)}{N} \right) \left((N - m) \tau^{(1)} + \mathbf{1}' \mathbf{u}^{(1)} \right) + \epsilon \pi, \\ \text{subject to} & \tau^{(k)} + \mathbf{u}_j^{(k)} \geq \gamma_j(\boldsymbol{\phi}) + 1, \quad k = 1, 2, \; j = 1, \ldots, N, \\ & \tau^{(k)} + \mathbf{u}_j^{(k)} \geq 0, \quad k = 1, 2, \; j = 1, \ldots, N, \\ & \pi \geq \kappa \gamma_j(\boldsymbol{\phi}) + 1, \quad j = 1, \ldots, N, \\ & \pi \geq 0, \mathbf{u} \geq 0, \boldsymbol{\psi} \in \tilde{\Phi}. \end{array} \tag{54}$$

DJIA	Dow Jones Composite Average
NDX	Nasdaq 100
RUT	Russell 2000

Table 2: Base set of factors

4 Computational Experiments

In this section we report the results of our numerical experiments with the portfolio selection model proposed in § 2. We conducted two sets of experiments. The first set of experiments compared the performance of a robust active portfolio management strategy with that of a non-robust version of that strategy on simulated market data. The results of this experiment are presented in § 4.1. The second set of experiments, described in § 4.2, compare the performance of the robust and non-robust strategies on real market data. In these experiments we do not use the alternative models described in §3.

In both sets of experiments we assumed that the transaction costs were given by (19) and (21), and that the parameters of the transaction cost functions $T_i(z_i, y_i)$ for all assets i were the same, i.e., $\vartheta i1 = \vartheta$ and $\pi_i = \pi$ for $i = 1, \ldots, n$. To set θ and π we relied on the data in [25], which was from 1982. That data suggests that $\vartheta = 0.01$ and the break point π corresponds to \$0.25mn for the large-cap stocks. We assumed that the parameter ϑ is the same today. However, since the price of the S&P 500 has increased approximately 10 times from 119.15 on August 3rd, 1982 to 1059.02 on November 3rd, 2003, we assumed that the break point π also increased 10 times. Hence, we set $\pi =$ \$2.5mn. We assumed our current wealth to be $w =$ \$100mn. We do not argue that these values model reality exactly – estimating the exact transaction cost function is beyond the scope of this paper. However, these transaction costs are reasonable in the light of the data in [25].

4.1 Experiments on simulated data

For this set of experiments the number of assets n was set to $n = 200$ and the number of factors m was set to $m = 38$. The linear model given by (1) and (3) implies that the market is defined by the factor covariance matrix $\mathbf{F}^s$, the factor loading matrix $\mathbf{V}^s$, the covariance of the residual returns $\mathbf{D}^s$, the "betas" of the assets $\boldsymbol{\beta}^s$, the mean return μ_b^s of the benchmark, the volatility σ^s of the benchmark and the mean exceptional return vector $\boldsymbol{\alpha}^s$. The superscript "**s**" indicates that these parameters were set at the beginning of the simulation and not randomly generated. To ensure that the simulated assets returns were realistic we used the following strategy to select these parameters. We randomly selected 35 stocks from the constituents of the S&P 500 on January 2, 2004 and the 3 major market indices in Table 2 and deemed them to be the factors driving the market. We set the factor covariance matrix $\mathbf{F}^s$ equal to the (38×38)-covariance matrix estimated using the past 300 days. Each element of the matrix $\mathbf{V}^s$ and the $\boldsymbol{\beta}^s$ was independently sampled from a

$\mathcal{N}(0, 0.5)$ distribution. We set the simulated benchmark's mean daily return $\mu_b^s = 6.5 \times 10^{-4}$ and the daily volatility $\sigma^s = 0.01$. Each diagonal element of $\mathbf{D}^s$ was sampled from a uniform distribution over the interval $[10^{-6}, 10^{-4}]$. The components of the vector $\boldsymbol{\alpha}^s$ were sampled from a $\mathcal{N}(0, 0.002)$ distribution. As noted above, the samples for the market parameters were generated once at the beginning of the simulation and were held fixed through the entire simulation run. Given the market parameters, the daily return $\mathbf{r}^{(t)} \in \mathbb{R}^n$ of the $n = 200$ assets is given by

$$\mathbf{r}^{(t)} = \mathcal{N}(\boldsymbol{\mu}^s, \boldsymbol{\Sigma}^s), \tag{55}$$

where

$$\boldsymbol{\mu}^s = \mu_b \boldsymbol{\beta}^s + \boldsymbol{\alpha}^s, \qquad \boldsymbol{\Sigma}^s = \sigma_b^2 \boldsymbol{\beta}^s (\boldsymbol{\beta}^s)' + (\mathbf{V}^s)'(\mathbf{F}^s)\mathbf{V}^s + \mathbf{D}^s$$

The details of the simulation experiment is described below.

(a) We re-balanced the portfolio every $T = 60$ trading dates. Thus, we bought and held portfolios for approximately a quarter. Our investment horizon was 9 periods.

(b) The investor only observes the market return data, i.e. in particular, the investor does not know the market parameters and in particular the factors driving the market. At the beginning of each investment period:

- We computed an estimate $\boldsymbol{\Sigma}$ of the covariance matrix of asset returns using the market data from the previous $H = 5$ periods, i.e. $H \times T = 300$ days.
- We assumed that the factors for the particular period were the three standard indices in Table 2, the benchmark and the eigenvectors corresponding to the largest eigenvalues of $\boldsymbol{\Sigma}$ accounting 95% of $\mathbf{Tr}(\boldsymbol{\Sigma})$. Note that this set of factors is *different* from the true set of factors driving the market.
- We estimated the vector $\boldsymbol{\beta}$ from the history of $H \times T$ days.
- We computed the estimate $\mathbf{F}$ of the factor covariance matrix using the simulated factor returns from the history of $H \times T$ days.
- We set the confidence level $\omega = 99\%$ and computed the parameters $\boldsymbol{\alpha}_0$, $\boldsymbol{\eta}$, $\mathbf{V}_0$, $\mathbf{G}$, $\boldsymbol{\rho}$, $\bar{\mathbf{d}}$, and $\boldsymbol{\delta}$ using the procedure described in Appendix A.

(c) We set the parameters $\theta = 0.2$, the upper bound $\mathbf{u} = 0.1\mathbf{1}$ and the lower bound $\mathbf{v} = 0.01\mathbf{1}$. See (18) for the definition of θ, $\mathbf{u}$ and $\mathbf{v}$.

(d) The non-robust portfolio $\boldsymbol{\phi}_o^{(p)}$ for period p was set equal to an optimal solution of

$$\begin{array}{ll} \text{maximize} & v_0(\boldsymbol{\phi}) \\ \text{subject to} & \boldsymbol{\phi} \in \Phi(\widehat{\boldsymbol{\phi}}_o^{(p)}), \end{array}$$

where $\widehat{\phi}_o^{(p)}(i) = (1+r_i^{(p-1)})\phi_o^{(p-1)}(i)$, $i = 1, \ldots, n$, $r_i^{(p-1)}$ is the return on asset i over the period p, $\Phi(\widehat{\boldsymbol{\phi}}_o^{(p)})$ denotes that the portfolio $\bar{\boldsymbol{\phi}}$ defining Φ (see 18) was set equal to $\widehat{\boldsymbol{\phi}}_o^{(p)}$.

Similarly, the robust active portfolio was set equal to an optimal solution of the robust problem

$$\begin{array}{ll} \text{maximize} & v_r(\boldsymbol{\phi}) \\ \text{subject to} & \boldsymbol{\phi} \in \Phi(\widehat{\boldsymbol{\phi}}_r^{(p)}), \end{array}$$

where $\widehat{\phi}_r^{(p)}(i) = (1+r_i^{(p-1)})\phi_r^{(p-1)}(i)$, $i = 1, \ldots, n$.

Both the non-robust active strategy and the robust active strategy started with an initial wealth $w^{(0)} = \$100$mn. In the first period we assumed that transaction costs were zero – this ensured that we did not have to specify an initial portfolio $\bar{\boldsymbol{\phi}}$ to compute $\boldsymbol{\phi}^{(1)}$.

(e) The results reported are the averages over 50 independent runs.

We summarize our numerical results in a series of plots highlighting different performance measures. Let $w^{(t)}$ denote the wealth of a portfolio strategy at the end of day t. Then

$$w^{(t)} = \sum_{i=1}^{n} \Big(\prod_{k=(p-1)T+1}^{t} (1 + r_f + r_i^{(k)}) \Big) \phi_i^{(p)}, \qquad (p-1)T < t \leq pT, \tag{56}$$

where the risk-free rate of return $r_f = 3\%$ per year, $r_i^{(k)}$ denote the return in excess of the risk-free rate r_f on asset i on day k, and $\boldsymbol{\phi}^{(p)}$ is the portfolio held by the portfolio strategy in period p. The relative daily wealth $rw^{(t)}$ of a strategy with respect to the index B at the end of day t is

$$rw^{(t)} = \frac{w^{(t)}}{w_B^{(t)}},$$

where $w_B^{(t)}$ denotes the cumulative wealth generated by investing the initial wealth $w^{(0)} = \$100$mn in the benchmark.

Figure 3 displays the relative wealth generated by the robust active strategy. In this plot the solid line is the relative wealth averaged over 50 simulation runs, the box at each period is the one standard deviation interval, and the vertical lines show the max and min values of the relative cumulative wealths in that period. For comparison, the average relative wealth generated by the non-robust strategy is depicted by a dotted line. Figure 4 displays the results for the non-robust strategy.

Figures 5 and 6 plot the per-period excess return R_p of the robust active strategy and the non-robust strategy respectively, where

$$R^{(p)} = \frac{w^{(Tp)}}{w^{(T(p-1))}} - \frac{w_B^{(Tp)}}{w_B^{(T(p-1))}},$$

i.e. $R^{(p)}$ is the return of the portfolio strategy in excess of the benchmark return.

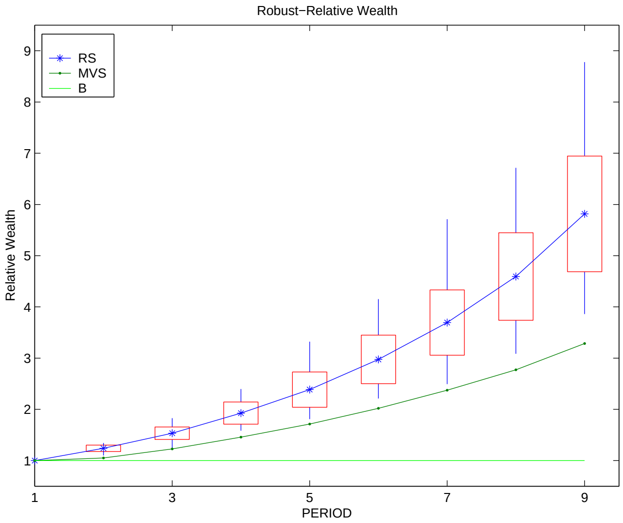

Figure 3: Relative wealth of the robust strategy

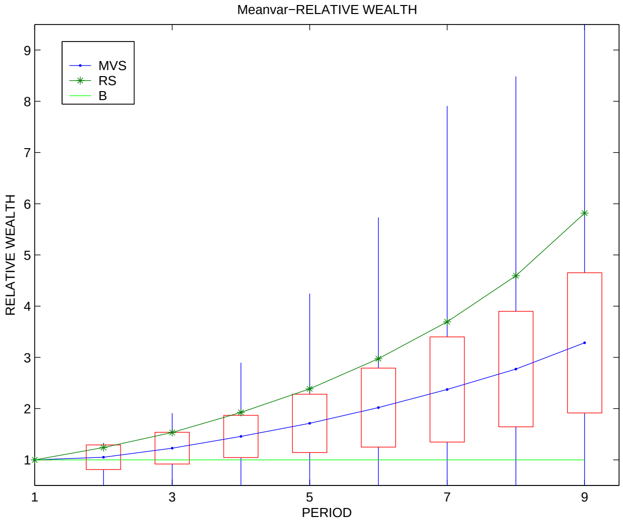

Figure 4: Relative wealth of the non-robust mean-variance strategy

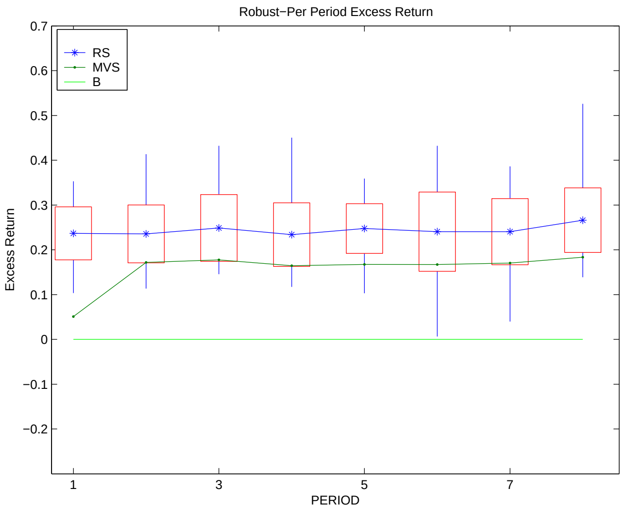

Figure 5: Per period excess return of the robust strategy

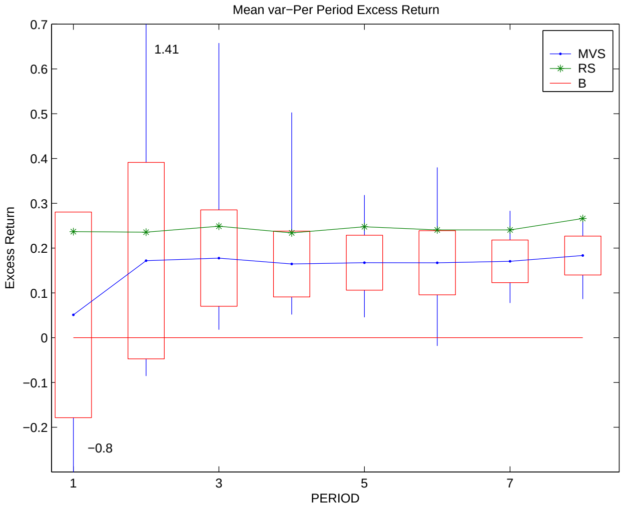

Figure 6: Per period excess return of the non-robust mean-variance strategy

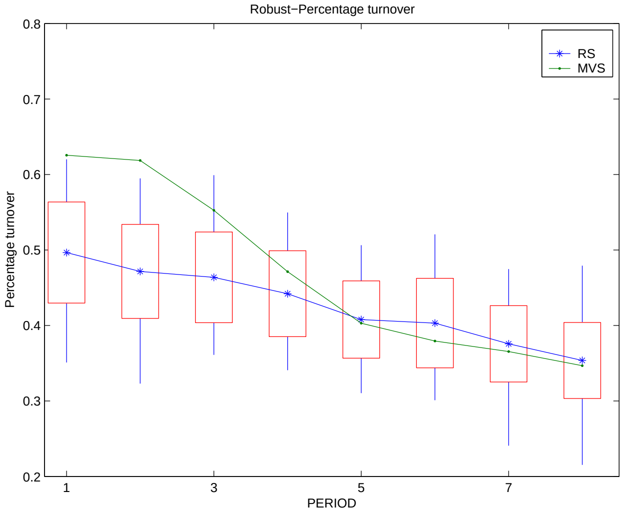

Figure 7: Turnover of the robust portfolios

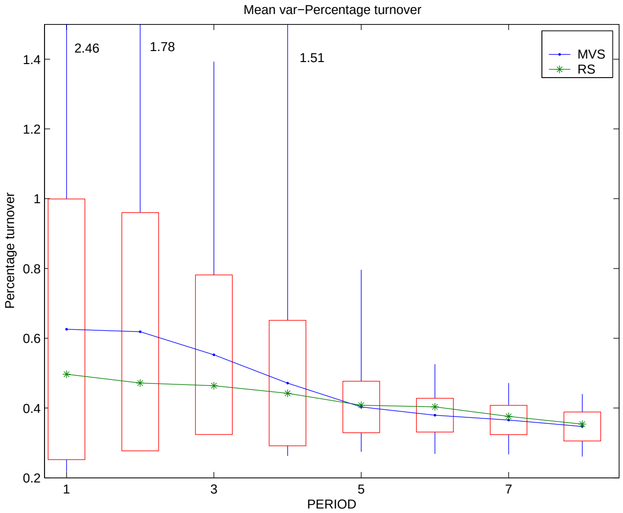

Figure 8: Turnover of the non-robust mean-variance portfolios

Figures 7 and 8 plot the portfolio turnover κ of, respectively, the robust and the non-robust strategies, where

$$\kappa = \frac{\sum |\boldsymbol{\phi}_i^{(p)} - \boldsymbol{\phi}_i^{(p-1)}|}{\sum |\boldsymbol{\phi}_i^{(p-1)}|}$$

Note that the percentage turnover can be greater than 1 because of short sales. The following observations are supported by the results displayed in Figures 3–8.

(i) The robust active portfolio strategy is always above the benchmark, i.e. it consistently provides exceptional returns. This is not true for the non-robust mean-variance active portfolio strategy – the minimum wealth of this strategy is often much below that of the benchmark in every period!

(ii) The *minimum* wealth of the robust portfolio strategy in any period is higher than that average wealth of the non-robust portfolio strategy.

(iii) The return of the robust portfolio strategy is more stable in the sense that ratio of the minimum to the maximum wealth in any period is larger than the same ratio for the non-robust strategy.

(iv) The turnover of the robust strategy and the non-robust strategy are roughly comparable on average, but the turnover of the non-robust strategy is more variable and can be quite large.

(v) The average number of assets in the robust portfolio was 109.4 and the number of assets in the classical mean-variance portfolio was 177.6. Therefore, on average, the total cost of managing the robust strategy ought to be lower.

Since the robust strategy is protecting against all parameters in the uncertainty set, one would expect that the robust strategy is likely to perform better in markets where the parameters fluctuate. The next set of experiments test this hypothesis by modifying the daily asset returns as follows:

$$\tilde{\mathbf{r}} = \begin{cases} \mathbf{r} = \mathcal{N}(\boldsymbol{\mu}^s, \boldsymbol{\Sigma}^s), & \text{with probability } 1 - \varsigma, \\ \bar{\mathbf{r}} = \mathcal{N}(\bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\Sigma}}), & \text{with probability } \varsigma, \end{cases}$$

where $\boldsymbol{\mu}^s = \boldsymbol{\beta}_i^s \mu_b s + \boldsymbol{\alpha}_i^s$ and $\boldsymbol{\Sigma}^s = \sigma_b^2 \boldsymbol{\beta}^s (\boldsymbol{\beta}^s)^T + (\mathbf{V}^s)'(\mathbf{F}^s)^{(}\mathbf{V}^s) + \mathbf{D}^s$ denote the mean and variance of the usual simulated returns, the shifted mean return

$$\bar{\mu}_i = \Big(1 - 0.1\text{sign}(\mu_i^s)\Big)\mu_i^s, \quad i = 1, \ldots, n,$$

and the shifted covariance matrix $\bar{\boldsymbol{\Sigma}} = \sigma_b^2 \boldsymbol{\beta}^s (\boldsymbol{\beta}^s)^T + (\bar{\mathbf{V}})'(\mathbf{F}^s)^{(}\bar{\mathbf{V}}) + \mathbf{D}$, with $\bar{V}_{ij} = v_{ij}(1 + z)$ and and $z \sim \mathcal{N}(0, 0.01)$. Note that in this new market model, with probability ς, the expected return $\bar{\mu}_i$ shifts in a direction opposite to that of the nominal return μ_i^s.

Figures 9 and 10 plot, respectively, the relative wealth of the robust and non-robust strategy for $\varsigma = 0.2$, i.e. when the returns deviate from the regular return model with 20% probability.

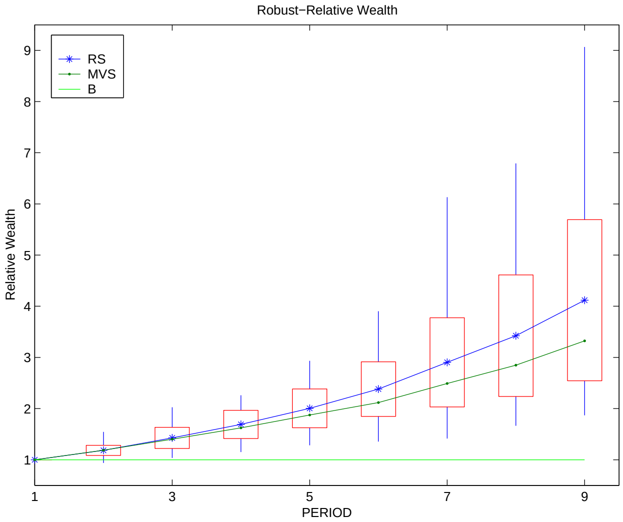

Figure 9: Relative wealth of the robust strategy when $\varsigma = 0.2$

Again, the results are averaged over 50 independent simulation runs. Figures 11 and 12 plot the excess return of the two portfolio strategies for $\varsigma = 0.2$ and Figures 13, and 14 display the turnover of the two portfolio strategies. When $\varsigma = 0.2$ the average number of assets held by the robust strategy and the non-robust strategies were 86.06 and 181.50, respectively. These plots support the following observations.

(i) The robust strategy outperforms the non-robust strategy; however, the excess returns of both strategies are lower than their respective returns in a stationary market.

(ii) The turnover of the robust strategy is significantly higher than the non-robust strategy. This fact, together with the observations that the robust strategy holds fewer assets and has a higher excess return, lends to the hypothesis that the robust strategy is better able to react to market shifts.

4.2 Real Market Data

In this section we report the results of computational experiments with real market data. The universe of assets in these set of experiments was the stocks that were part of the S&P 500 index and the benchmark was the S&P 500 index. Thus, the goal of the active strategies was to provide returns in excess of the S&P 500 index. The strategies invested over the period April 1, 2001 to November 10, 2003 using data starting from January 5, 2000.

The experimental procedure was very similar to the one described in the previous section. However, in contrast to the simulation runs, here we have a single sample path.

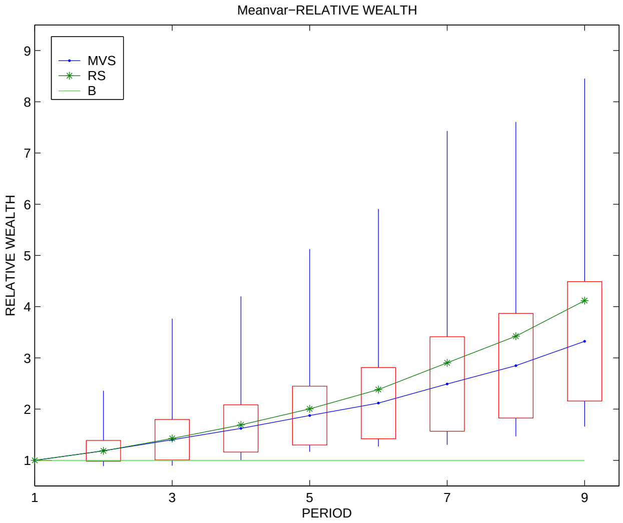

Figure 10: Relative wealth of the non-robust strategy when $\varsigma = 0.2$

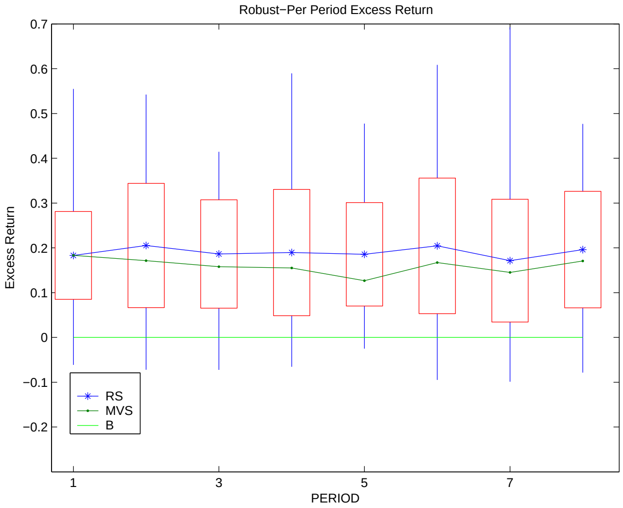

Figure 11: Per period excess return of the robust strategy when $\varsigma = 0.2$

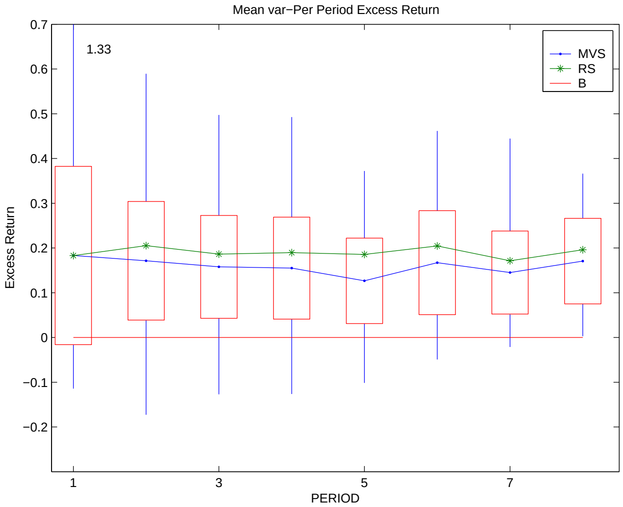

Figure 12: Per period excess return of the non-robust strategy when $\varsigma = 0.2$

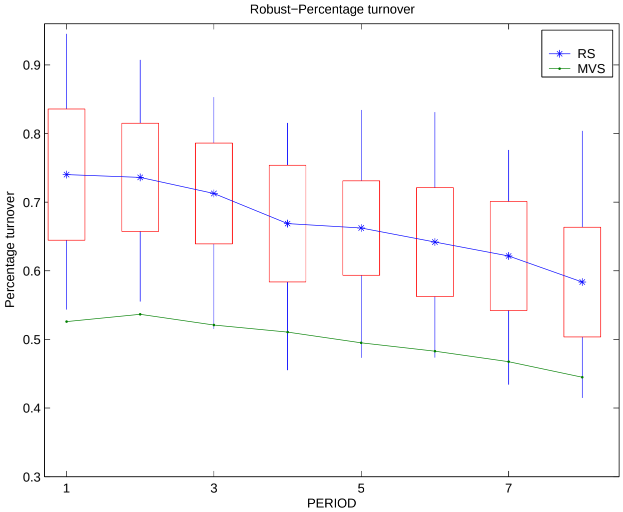

Figure 13: Turnover of the robust portfolios when $\varsigma = 0.2$

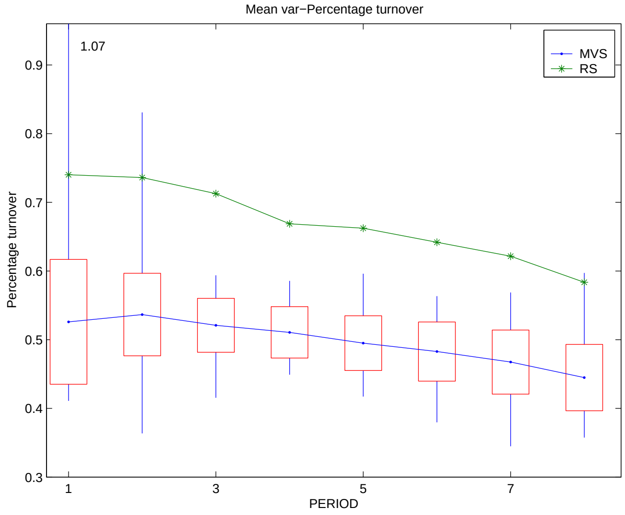

Figure 14: Turnover of the non-robust portfolios when $\varsigma = 0.2$

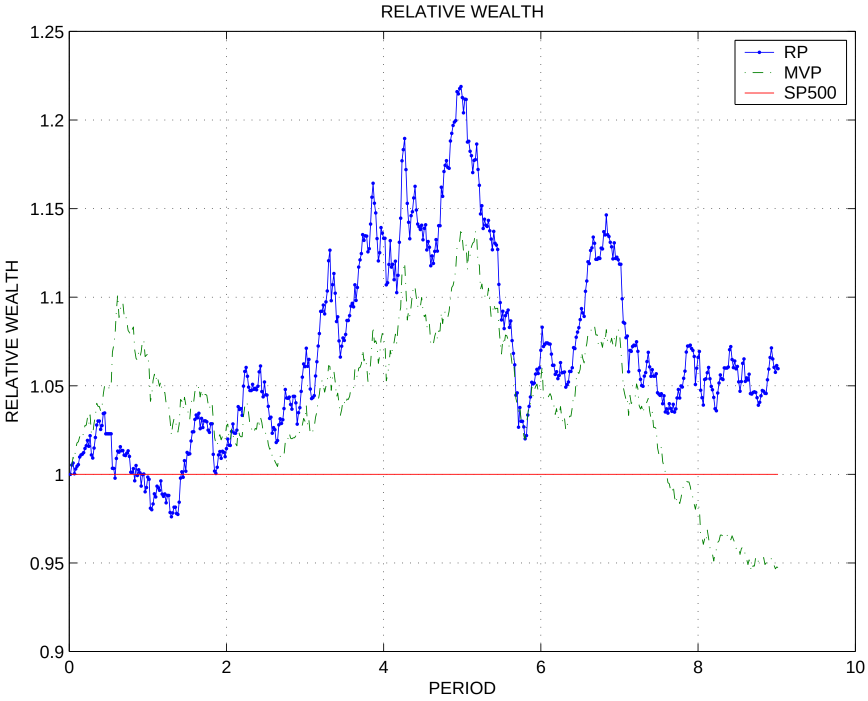

Figure 15: Relative wealth of the strategies

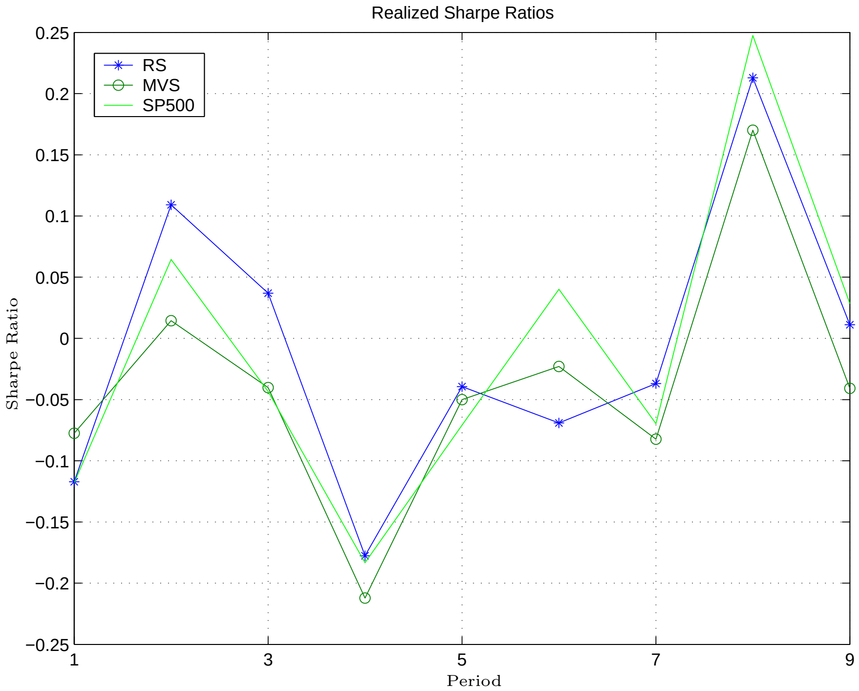

Figure 16: Realized sharpe ratios of the strategies

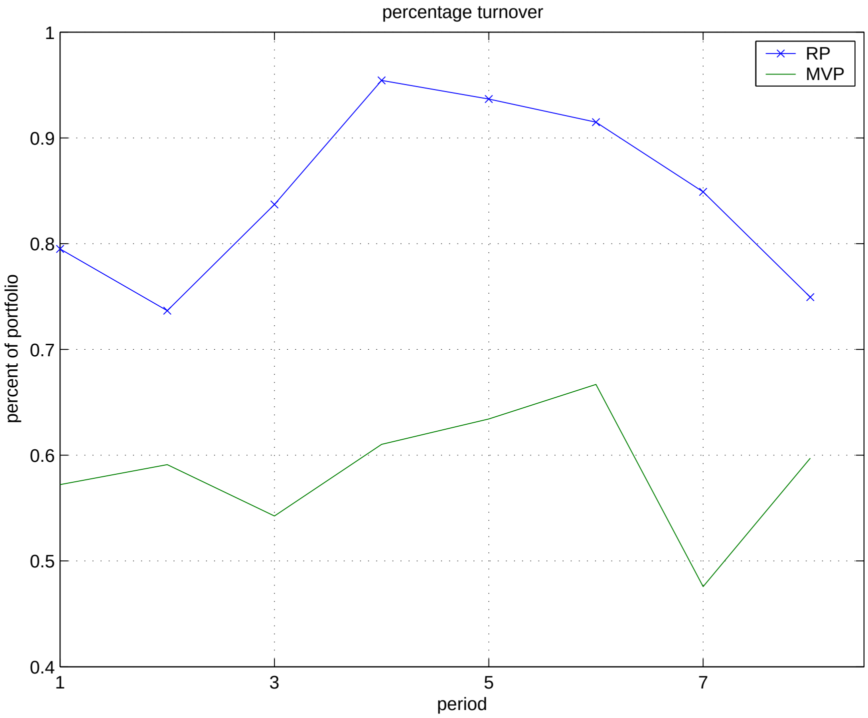

Figure 17: Percentage turnover of the portfolios

Figure 15 plots the relative wealth of the strategies with respect to that of S&P 500 and Figure 17 shows the percentage turnover of the portfolios. Figure 16 displays the realized Sharpe ratio of the two portfolio strategies and the S&P 500 index. The average number of assets held in the portfolios were 72.2 for the robust strategy and 413.4 for the non-robust strategy. Thus, the non-robust strategy hold pretty much all the stocks in the index.

Since the results above refer to a single sample path, there is a chance that they suffer from a starting-point bias. The next set of results attempt to correct this bias. We considered $N = 60$ different strategies that each re-balance every $T = 60$ trading days but each have different start date. Thus, a strategy that starts investing on day d, $d = 0, \ldots, N-1$, re-balances the portfolio on day $60k + d$, $k \geq 0$. Let $w^{(t)}(d)$ denote the wealth on day t, $t \geq 1$, of the portfolio strategy that starts investing on day d and let $w_B^t(d)$ denote the cumulative total wealth on day t when the initial investment of \$100mn is invested in the S&P 500 index on day d. If the S&P 500 index has a stationary distribution. Then the sequences of relative wealth

$$rw^{(t)}(d) = \frac{w^{(t+d)}(d)}{w_B^{(t+d)}(d)}, \quad t \geq 0,$$

corresponding to each starting date d are independent and identically distributed. Therefore, one can treat these different sequences as independent investment runs. In Figure 18 the solid lines correspond to the average

$$\frac{1}{N} \sum_{d=0}^{N-1} rw^{(t)}(d)$$

for each of the two strategies, and the box at each period is the one standard deviation interval for the relative wealth of the robust strategy, and the vertical lines show the max and min values of the relative wealths in that period for the robust strategy. Figure 19 displays the analogous quantities for the non-robust strategy. From these plots it is clear that the robust strategy is superior to the non-robust strategy and the insights from the simulation model continue to hold for the real market data.

5 Conclusion

In this paper we show how to use robust optimization based techniques to immunize the active portfolio selection problems to perturbations in parameter values. In §2 and §3 we discuss a number of different models for active portfolio management and show to make them robust to data perturbations. We show that the portfolio selection problem in all of these models can be reformulated as an SOCP. We also develop a piece-wise convex model for trading costs which allows a great deal of modeling flexibility without increasing the computational cost of solving the portfolio selection problem – they still remain SOCPs. This fact has important theoretical and practical implications. Since the computational complexity of an SOCP is comparable to that

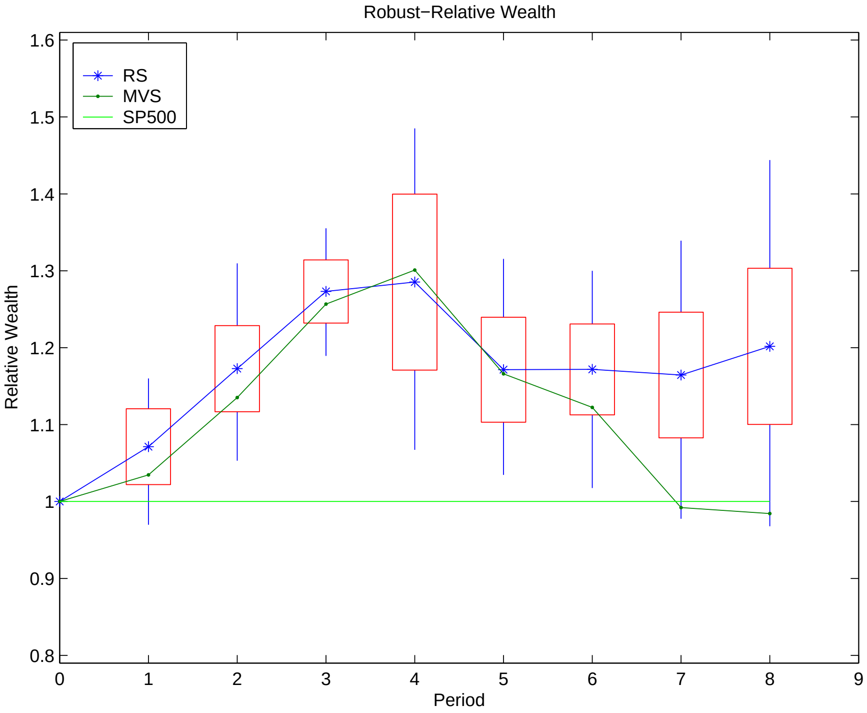

Figure 18: Relative cumulative wealth of the robust strategy with respect to the S&P 500

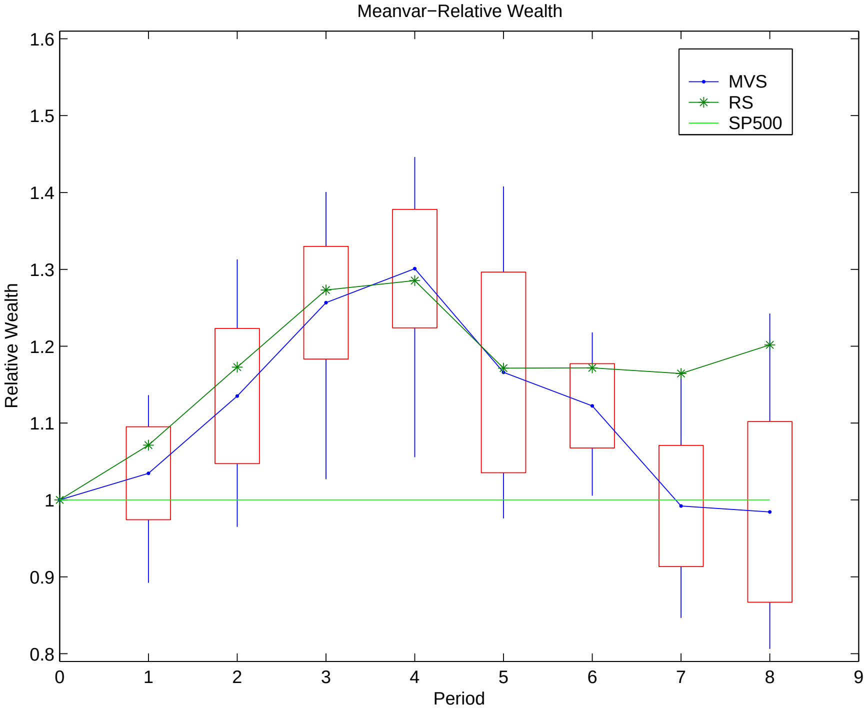

Figure 19: Relative cumulative wealth of the non-robust strategy with respect to the S&P 500

of a convex quadratic program, it follows that robust active portfolio selection is able to provide protection against parameter fluctuations at very moderate computational cost. Morever, a number commercial solvers such as MOSEK, CPLEX and Frontline System (supplier of EXCEL SOLVER) provide the capability for solving SOCPs in a numerically robust manner.

There are several interesting extensions that are worth exploring. The models in this paper are all single-period myopic models. Extending the portfolio selection model to a multi-period setting appears to be non-trivial.

References

[1] F. Alizadeh and Goldfarb D. Second-order cone programming. *Mathematical Programming*, 95(1):3–51, 2003.

[2] T. W. Anderson. *An Introduction to Multivariate Statistical Analysis.* John Wiley & Sons, 2nd edition, 1984.

[3] A. Ben-Tal and A. Nemirovski. Robust convex optimization. *Math. Oper. Res.*, 23(4):769–805, 1998.

[4] A. Ben-Tal and A. Nemirovski. Robust solutions of uncertain linear programs. *Oper. Res. Lett.*, 25(1):1–13, 1999.

[5] A. Ben-Tal and A. Nemirovski. *Lectures on modern convex optimization.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2001. Analysis, algorithms, and engineering applications.

[6] D. Bertsimas and D. B. Brown. Robust linear optimization and coherent risk measures. Technical Report 2659, LIDS, MIT, 2005.

[7] F. Black and R. Litterman. Asset allocation: combining investor views with market equilibrium. *Journal of Fixed Income*, pages 7–18, September 1991.

[8] V. K. Chopra. Improving optimization. *J. Investing*, 2(Fall):51–59, 1993.

[9] V. K. Chopra, C. R. Hensel, and A. L. Turner. Massaging mean-variance inputs: Returns from alternative investment strategies in the 1980s. *Manag. Sci.*, 39(7):845–855, Jul 1993.

[10] J. Cvitanic, A. Lazrak, L. Martellini, and F Zapatero. Revisiting Treynor and Black (1973): a second step towards a theory of active portfolio Management. Second revision, January 2003.

[11] B. Efron and R. Tibshirani. *An Introduction to Bootstrap.* Number 57 in Monographs on Statistics and Applied Probability. Chapman & Hall, New York, 1993.

[12] E. J. Elton and M. J. Gruber. *Modern portfolio theory and investment analysis.* John Wiley & Sons, INC., fifth edition, 1995.

[13] P. A. Frost and J. E. Savarino. An empirical Bayes approach to efficient portfolio selection. *J. Fin. Quan. Anal.*, 21:293–305, 1986.

[14] P. A. Frost and J. E. Savarino. For better performance: constrain portfolio weights. *J. Portfolio Manag.*, 15:29–34, 1988.

[15] D. Goldfarb and G. Iyengar. Robust portfolio selection problems. *Mathematics of Operations Research*, 28(1):1–38, 2003.

[16] W. H. Greene. *Econometric analysis.* Macmillan Pub. Co., 1990.

[17] R. C. Grinold and R. N. Kahn. *Active Portfolio Management.* McGraw-Hill Professional Publishing, 2nd edition edition, 1999.

[18] L. P. Hansen and T. J. Sargent. Robust control and model uncertainty. *American Economic Review*, 91:60–66, 2001.

[19] G. He and R. Litrerman. The intuition behind Black-Litterman model portfolios. Technical report, Goldman, Sachs & Co., Fixed Income Research, December 1999.

[20] G. Iyengar. A note on the sample approximation of cvar. Technical Report, IEOR Department, 2006.

[21] R. W. Klein and V. S. Bawa. The effect of estimation risk on optimal portfolio choice. *J. Finan. Econ.*, 3:215–231, 1976.

[22] L. Kogan and T. Wang. A simple theory of asset pricing under model uncertainty. Working paper, University of British Columbia, 2002.

[23] J. Lintner. Valuation of risky assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics*, 47:13–37, 1965.

[24] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret. Applications of second-order cone programming. *Linear Algebra Appl.*, 284(1-3):193–228, 1998.

[25] T. F. Loeb. Trading cost: The critical link between investment information and results. *Financial Analysts Journal*, May-June:39–44, 1983.

[26] H. M. Markowitz. Portfolio selection. *J. Finance*, 7:77–91, 1952.

[27] H. M. Markowitz. *Portfolio Selection.* Wiley, New York, 1959.

[28] R. O. Michaud. *Efficient Asset Management: a practical guide to stock portfolio optimization and asset allocation.* Harward Business School Press, Boston, 1998.

[29] J. Mossin. Equilibrium in capital asset markets. *Econometrica*, 34(4):768–783, 1966.

[30] A. Nemirovski and A. Shapiro. Convex approximation of chance constrained problems. To appear in *SIAM J. Opt.* Available at `http://www.optimization-online.org/DB_HTML/2004/12/1033.html`, 2004.

[31] Y. Nesterov and A. Nemirovski. *Interior-point polynomial algorithms in convex programming.* SIAM, Philadelphia, 1993.

[32] L. Pastor and R. F. Stambaugh. Comparing asset-pricing models: an investment perspective. *Journal of Financial Economics*, 56:335–381, 2000.

[33] N.G. Polson and B. V. Tew. Bayesian portfolio selection: an empirical analysis of the s&p 500 index 1970-1996. *Journal of Business and Economics Statistics*, 18:164–173, 2000.

[34] R. T. Rockafellar and S. Uryasev. Optimization of conditional value-at-risk. *Journal of Risk*, 2(3):21–41, 2000.

[35] W. Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, 19(3):425–442, 1964.

[36] J. Stürm. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software*, 11–12:625–653, 1999.

[37] J.L. Treynor and F. Black. How to use security analysis to improve portfolio selection. *Journal of Business*, pages 66–88, 1973.

[38] R. Tütüncü and M. Koenig. Robust asset allocation. *Annals of Operations Research*, 132:157–187, 2004.

[39] Z. Wang. A shrinkage approach to model uncertainty and asset allocation. forthcoming in *The Review of Financial Studies*, 2004.

[40] W. T. Ziemba and J. M Mulvey, editors. *Worldwide Asset and Liability Modeling.* Cambridge Univ. Press, Cambridge, UK, 1998.

A Defining the uncertainty sets

Suppose the market data consists of asset excess returns $\{\mathbf{r}^{(t)} : t = 1, ..., T\}$, benchmark returns $\{r_b^{(t)} : t = 1, ..., T\}$, and corresponding factor returns $\{\mathbf{f}^{(t)} : t = 1, ..., T\}$ for T trading days. After computing $\boldsymbol{\beta}$ by linear regression and forming $\Delta \mathbf{r}_i^{(t)} = \mathbf{r}_i^{(t)} - \boldsymbol{\beta}_i r_b^{(t)}$ for $i = 1, \ldots, n$ and $t = 1, \ldots, T$, the linear model (3) implies that

$$\Delta r_i^{(t)} = \alpha_i + \sum_{j=1}^{n} V_{ji} f_j^{(t)} + \epsilon_i^{(t)}, \quad i = 1, \ldots, n, \quad t = 1, \ldots, T. \tag{57}$$

In linear regression analysis, typically it is assumed that $\{\epsilon_i^{(t)} : i = 1, \ldots, n, t = 1, \ldots, T\}$ are all independent normal variables and $\epsilon_i^{(t)} \sim \mathcal{N}(0, \sigma_i^2\}$, for all $t = 1, \ldots, T$. The independence assumption is relaxed in ARMA models at the cost of replacing the least-squares estimation by Kalman filters (see [18]).

Let $\mathbf{B} = [\mathbf{f}^1, \mathbf{f}^2, \ldots, \mathbf{f}^T] \in \mathbf{R}^{m \times (T)}$ be the matrix of factor returns. Collecting together terms corresponding to a particular asset i over all periods $t = 1, \ldots, T$, we get the following linear model for the returns $\{r_i^{(t)} : t = 1, \ldots, T\}$,

$$\mathbf{y}_i = \mathbf{A}\mathbf{x}_i + \boldsymbol{\epsilon}_i,$$

where

$$\mathbf{y}_i = [\ \Delta r_i^1 \quad \Delta r_i^2 \quad \cdots \quad \Delta r_i^T\ ]', \quad \mathbf{A} = [\ \mathbf{1} \quad \mathbf{B}'\ ], \quad \mathbf{x}_i = [\ \alpha_i \quad V_{1i} \quad \cdots \quad V_{mi}\ ]'$$

and $\boldsymbol{\epsilon}_i = [\epsilon_i^{(1)}, \ldots, \epsilon_i^{(TH)}]'$ is the vector of residual returns corresponding to asset i. The least-squares estimate $\bar{\mathbf{x}}_i$ of the true parameter $\mathbf{x}_i$ is given by

$$\bar{\mathbf{x}}_i = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}' y_i,$$

if $\mathbf{rank}(\mathbf{A}) = m + 1$. Substituting $\mathbf{y}_i = \mathbf{A}\mathbf{x}_i + \boldsymbol{\epsilon}_i$, we get

$$\bar{\mathbf{x}}_i - \mathbf{x}_i = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\boldsymbol{\epsilon}_i \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}),$$

where $\boldsymbol{\Sigma} = \sigma_i^2(\mathbf{A}'\mathbf{A})^{-1}$. Let $\mathbf{Q} \in \mathbf{R}^{J \times (m+1)}$. A standard result in regression theory states that if σ_i^2 in the definition of $\boldsymbol{\Sigma}$ is replaced by s_i^2, where s_i^2 is the unbiased estimate of σ_i^2 and given by

$$s_i^2 = \frac{\|\mathbf{y}_i - \mathbf{A}\bar{\mathbf{x}}_i\|^2}{T - m - 1},$$

then

$$\nu = \frac{1}{J s_i^2}(\mathbf{Q}\bar{\mathbf{x}}_i - \mathbf{Q}\mathbf{x}_i)'(\mathbf{Q}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{Q}')^{-1}(\mathbf{Q}\bar{\mathbf{x}}_i - \mathbf{Q}\mathbf{x}_i)$$

is distributed according to the F-distribution with J degrees of freedom in the numerator and $T - m - 1$ degrees of freedom in the denominator [2, 16]; i.e., the probability $\nu \leq c_J(\omega)$ is ω, or equivalently,

$$\mathbf{P}\Big((\mathbf{Q}\bar{\mathbf{x}}_i - \mathbf{Q}\mathbf{x}_i)'(\mathbf{Q}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{Q}')^{-1}(\mathbf{Q}\bar{\mathbf{x}}_i - \mathbf{Q}\mathbf{x}_i) \leq J c_J(\omega) s_i^2\Big) = \omega \tag{58}$$

If we define $\mathbf{Q} = [\mathbf{e}_2, \mathbf{e}_3, \ldots, \mathbf{e}_{m+1}] \in \mathbf{R}^{m\times(m+1)}$, then $\mathbf{Q}\bar{\mathbf{x}}_i = \bar{\mathbf{V}}_i$ is the least squares estimate to the true factor loading $\mathbf{Q}\mathbf{x}_i = \mathbf{V}_i$ and using (58) we get

$$\mathbf{P}\Big((\bar{\mathbf{V}}_i - \mathbf{V}_i)'(\mathbf{Q}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{Q}')^{-1}(\bar{\mathbf{V}}_i - \mathbf{V}_i) \leq mc_m(\omega)s_i^2\Big) = \omega \tag{59}$$

Therefore, the set $S_v(\omega)$

$$S_v(\omega) = \Big\{\mathbf{V} : \mathbf{V} = \mathbf{V}_0 + \mathbf{W}, \|\mathbf{W}_i\|_g \leq \rho_i, i = 1, ..., n\Big\},$$

where

$$\mathbf{V}_0 = \bar{\mathbf{V}}, \quad \mathbf{G} = (\mathbf{Q}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{Q}')^{-1} = \mathbf{B}\mathbf{B}' - \frac{1}{TH}(\mathbf{B}\mathbf{1})(\mathbf{B}\mathbf{1})', \quad \rho_i = \sqrt{mc_m(\omega)s_i^2}, \quad i = 1, \ldots, n,$$

is an ω^n-confidence set for the factor loading matrix $\mathbf{V}$.

To construct the uncertainty set for expected residual return, define $\mathbf{Q} = \mathbf{e}_1'$. Then, $\mathbf{Q}\mathbf{x}_i = \alpha_i$ and $\mathbf{Q}\bar{\mathbf{x}}_i = \bar{\alpha}_i$ are the true expected residual return of asset i and the least squares estimate of the expected residual return, respectively. Therefore, (58) implies

$$\mathbf{P}\Big(|\bar{\alpha}_i - \alpha_i| \leq \sqrt{(\mathbf{A}'\mathbf{A})^{-1}_{11} c_1(\omega)s_i^2}\Big) = \omega \tag{60}$$

Since the residual returns $\boldsymbol{\epsilon}_i$ are assumed to be independent, it follows that

$$S_\alpha(\omega) = \{\boldsymbol{\alpha} : \boldsymbol{\alpha} = \boldsymbol{\alpha}_0 + \boldsymbol{\xi}, |\xi_i| \leq \eta_i, i = 1, ..., n\}, \tag{61}$$

where

$$\alpha_{0,i} = \bar{\alpha}_i, \quad \eta_i = \sqrt{(\mathbf{A}'\mathbf{A})^{-1}_{11} c_1(\omega)s_i^2}, \quad i = 1, ...n,$$

is an ω^n confidence region for the mean return $\boldsymbol{\mu}$. Observing

$$\begin{aligned}\mathbf{P}((\boldsymbol{\mu}, \mathbf{V}) \in S_\alpha(\omega) \times S_v(\omega)) &= 1 - \mathbf{P}((\boldsymbol{\mu}, \mathbf{V}) \notin S_\alpha(\omega) \times S_v(\omega)), \\ &\geq 1 - \mathbf{P}((\boldsymbol{\mu}) \notin S_\alpha(\omega) - \mathbf{P}(\mathbf{V} \notin S_v(\omega)), \\ &= 2\omega^n - 1,\end{aligned}$$

we conclude that the Cartesian product $S(\omega) = S_\alpha(\omega) \times S_v(\omega)$ is a joint confidence region of $(\boldsymbol{\alpha}, \mathbf{V})$ with $(2\omega^n - 1)$ confidence.

One way of constructing the set S_d, is to use a bootstrap ω-confidence interval [11] around each σ_i^2. In our formulation we used the mean residual variance $\overline{d}_i$, for each i, for simplicity.

B Convex relaxation

Suppose $\boldsymbol{\phi} \in \Phi$, where Φ is defined in (18) and $\mathbf{1}'\boldsymbol{\phi} + T(\mathbf{z}, \mathbf{y}) < \mathbf{1}'\bar{\boldsymbol{\phi}}$. We show below that there exists a portfolio $\widehat{\boldsymbol{\phi}} \in \Phi$ that satisfies

$$\mathbf{1}'\widehat{\boldsymbol{\phi}} + T(\widehat{\mathbf{z}}, \widehat{\mathbf{y}}) = \mathbf{1}'\bar{\boldsymbol{\phi}}. \tag{62}$$

with $v_o(\boldsymbol{\phi}) = v_o(\widehat{\boldsymbol{\phi}})$ and $v_r(\boldsymbol{\phi}) = v_r(\widehat{\boldsymbol{\phi}})$. Hence, relaxing the constraint (62) does not result in any reduction in the optimal objective values of either the robust or non-robust active portfolio selection problems.

Define a new solution $(\widehat{\boldsymbol{\phi}}, \widehat{\mathbf{z}}, \widehat{\mathbf{y}}, \widehat{w})$ as follows.

$$\begin{aligned} \widehat{\boldsymbol{\phi}} &= s\boldsymbol{\phi}, \\ \widehat{\mathbf{z}} &= \mathbf{z} + (s-1)\max\{\boldsymbol{\phi}, \mathbf{0}\}, \\ \widehat{\mathbf{y}} &= \mathbf{y} + (s-1)\max\{-\boldsymbol{\phi}, \mathbf{0}\}, \\ \widehat{w} &= sw, \end{aligned} \tag{63}$$

where the max operator is taken component wise. Then

$$\begin{aligned} \mathbf{1}'\widehat{\boldsymbol{\phi}} + T(\widehat{\mathbf{z}}, \widehat{\mathbf{y}}) - (\mathbf{1}'\boldsymbol{\phi} + T(\mathbf{z}, \mathbf{y})) &\geq (s-1)\mathbf{1}'\boldsymbol{\phi} + [\boldsymbol{\rho}_z' \ \boldsymbol{\rho}_y'] \begin{pmatrix} \widehat{\mathbf{z}} - \mathbf{z} \\ \widehat{\mathbf{y}} - \mathbf{y} \end{pmatrix} \\ &= (s-1)[w + (\boldsymbol{\rho}_z + \boldsymbol{\rho}_y)'|\boldsymbol{\phi}|] \end{aligned} \tag{64}$$

where $\boldsymbol{\rho} = (\boldsymbol{\rho}_z, \boldsymbol{\rho}_y)$ is a subgradient of $T(\cdot, \cdot)$ at $(\mathbf{z}, \mathbf{y})$. Since $T(\mathbf{z}, \mathbf{y})$ is an increasing function, $\boldsymbol{\rho} > \mathbf{0}$ and $w + (\boldsymbol{\rho}_z + \boldsymbol{\rho}_y)'|\boldsymbol{\phi}| > 0$. Thus, we can choose an $s > 1$ such that $(\widehat{\boldsymbol{\phi}}, \widehat{\mathbf{z}}, \widehat{\mathbf{y}})$ satisfies (62).

We will now show that $(\widehat{\boldsymbol{\phi}}, \widehat{\mathbf{z}}, \widehat{\mathbf{y}}, \widehat{w})$ satisfies the remaining constraints defining $\mathcal{S}$. First note that

$$\begin{aligned} \widehat{\boldsymbol{\phi}} - \widehat{\mathbf{z}} + \widehat{\mathbf{y}} &= s\boldsymbol{\phi} - \mathbf{z} - (s-1)\max\{\boldsymbol{\phi}, \mathbf{0}\} + \mathbf{y} + (s-1)\max\{-\boldsymbol{\phi}, \mathbf{0}\} \\ &= s\boldsymbol{\phi} - (s-1)\boldsymbol{\phi} - \mathbf{z} + \mathbf{y} \\ &= \boldsymbol{\phi} - \mathbf{z} + \mathbf{y} \quad = \quad \bar{\boldsymbol{\phi}}. \end{aligned}$$

Also,

$$\mathbf{1}'(\widehat{\mathbf{y}} - \widehat{\mathbf{z}}) = \mathbf{1}'(\mathbf{y} - \mathbf{z}) - (s-1)\mathbf{1}'\boldsymbol{\phi} \leq \theta w - (s-1)w < \theta w < \theta \widehat{w}.$$

Clearly $\widehat{\mathbf{z}} \geq \mathbf{0}, \widehat{\mathbf{y}} \geq \mathbf{0}$, and $\widehat{w} \geq 0$, and all the remaining constraints defining Φ are satisfied because they are all homogeneous of degree 1 in $\boldsymbol{\phi}$ and w.

Finally, since v_0 and v_r are scale-invariant, it follows that $v_o(\boldsymbol{\phi}) = v_o(\widehat{\boldsymbol{\phi}})$ and $v_r(\boldsymbol{\phi}) = v_r(\widehat{\boldsymbol{\phi}})$.